

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using fw model

Run on: January 31, 2003, 07:42:31 (Search time 73 Seconds
(without alignments)
1643.201 Million cell updates/sec

Title: US-09-807-201-7
Sequence: 1 GAGCATCCGATACATC.....CGACGCCAATCCCATCC 267

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	267	100.0	US-10-012-896-896	Sequence 896, App
2	267	100.0	US-09-895-793-896	Sequence 896, App
3	267	100.0	US-09-895-814-896	Sequence 896, App
4	267	100.0	US-09-759-143-896	Sequence 896, App
5	267	100.0	US-09-780-669-896	Sequence 896, App
6	267	100.0	US-09-822-827-896	Sequence 896, App
7	267	100.0	US-10-012-896-894	Sequence 894, App
8	267	100.0	US-09-895-793-894	Sequence 894, App
9	267	100.0	US-09-895-814-894	Sequence 894, App
10	267	100.0	US-09-759-143-894	Sequence 894, App
11	267	100.0	US-09-780-669-894	Sequence 894, App
12	267	100.0	US-09-822-827-894	Sequence 894, App
13	267	100.0	US-10-012-896-929	Sequence 929, App
14	267	100.0	US-09-895-793-929	Sequence 929, App
15	267	100.0	US-09-895-814-929	Sequence 929, App
16	267	100.0	US-09-759-143-929	Sequence 929, App
17	267	100.0	US-09-780-669-929	Sequence 929, App
18	267	100.0	US-09-822-827-929	Sequence 929, App
19	265.4	99.4	US-09-981-353-922	Sequence 22, App

20	219	82.0	1476	9	US-10-012-896-931	Sequence 931, App
21	219	82.0	1476	9	US-09-895-793-931	Sequence 931, App
22	219	82.0	1476	9	US-09-895-814-931	Sequence 931, App
23	219	82.0	1476	10	US-09-759-143-931	Sequence 931, App
24	219	82.0	1476	10	US-09-780-669-931	Sequence 931, App
25	219	82.0	1476	10	US-09-822-827-931	Sequence 931, App
26	219	82.0	1479	9	US-10-012-896-930	Sequence 930, App
27	219	82.0	1479	9	US-09-895-793-930	Sequence 930, App
28	219	82.0	1479	9	US-09-895-814-930	Sequence 930, App
29	219	82.0	1479	10	US-09-759-143-930	Sequence 930, App
30	219	82.0	1479	10	US-09-780-669-930	Sequence 930, App
31	219	82.0	1479	10	US-09-822-827-930	Sequence 930, App
32	178	66.7	196	9	US-09-232-880-151	Sequence 151, App
33	178	66.7	196	9	US-10-012-896-151	Sequence 151, App
34	178	66.7	196	9	US-09-895-793-151	Sequence 151, App
35	178	66.7	196	9	US-09-895-814-151	Sequence 151, App
36	178	66.7	196	10	US-09-759-143-151	Sequence 151, App
37	178	66.7	196	10	US-09-780-669-151	Sequence 151, App
38	178	66.7	196	10	US-09-822-827-151	Sequence 151, App
39	178	66.7	196	10	US-09-822-827-151	Sequence 151, App
40	178	66.7	196	10	US-09-115-453-151	Sequence 151, App
41	32.2	12.1	369	10	US-09-860-352-4371	Sequence 4371, App
42	32.2	12.1	412	10	US-09-864-761-22443	Sequence 22443, App
43	32.2	12.1	442	10	US-09-864-761-5680	Sequence 5680, App
44	32	12.0	2479	9	US-10-012-896-894	Sequence 894, App
45	32	12.0	2479	9	US-09-895-793-894	Sequence 894, App

ALIGNMENTS

RESULT 1
US-10-012-896-896
Sequence 896, Application US/10012896
Publication No. US20020183251A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kaios, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasar A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Foy, Gary R.
APPLICANT: Mantanabe, Yoshitiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012, 896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 896
LENGTH: 683
TYPE: DNA
ORGANISM: Homo sapiens
US-10-012-896-896

Query Match 100.0%; Score 267; DB 9; Length 683;
Best Local Similarity 100.0%; Pred. No. 5.7e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATTAATGATGCTGTGATTAACAAGATGGCTTTGAAC 60
DB 9 GAACATTCAGATACCTATCATTAATGATGCTGTGATTAACAAGATGGCTTTGAAC 68
QY 61 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
DB 69 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
QY 121 CCTATTCGCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGCAATCCGCTCAGTAC 180
DB 129 CCTATTCGCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGCAATCCGCTCAGTAC 188
QY 181 TACCCGTCCCGGTGCGCCGAGTACGCCCGAGGGTCTCTGACGAGGTTCCAAACCCCGTC 240
DB 189 TACCCGTCCCGGTGCGCCGAGTACGCCCGAGGGTCTCTGACGAGGTTCCAAACCCCGTC 248
QY 241 GTCTGACGACGACCCCAATATCCCATCC 267
DB 249 GTCTGACGACGACCCCAATATCCCATCC 275

RESULT 2

US-09-895-793-896
Sequence 896, Application US/09895793
Publication No. US20020192763A1
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yudi
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.534C2
CURRENT APPLICATION NUMBER: US/09/895,793
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 896
LENGTH: 683
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-793-896

Query Match 100.0%; Score 267; DB 9; Length 683;
Best Local Similarity 100.0%; Pred. No. 5.7e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATTAATGATGCTGTGATTAACAAGATGGCTTTGAAC 60
DB 9 GAACATTCAGATACCTATCATTAATGATGCTGTGATTAACAAGATGGCTTTGAAC 68

QY 61 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
DB 69 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
QY 121 CCTATTCGCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGCAATCCGCTCAGTAC 180
DB 129 CCTATTCGCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGCAATCCGCTCAGTAC 188
QY 181 TACCCGTCCCGGTGCGCCGAGTACGCCCGAGGGTCTCTGACGAGGTTCCAAACCCCGTC 240
DB 189 TACCCGTCCCGGTGCGCCGAGTACGCCCGAGGGTCTCTGACGAGGTTCCAAACCCCGTC 248
QY 241 GTCTGACGACGACCCCAATATCCCATCC 267
DB 249 GTCTGACGACGACCCCAATATCCCATCC 275

RESULT 3

US-09-895-814-896
Sequence 896, Application US/09895814
Publication No. US20020193296A1
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yudi
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C6
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 896
LENGTH: 683
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-814-896

Query Match 100.0%; Score 267; DB 9; Length 683;
Best Local Similarity 100.0%; Pred. No. 5.7e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATTAATGATGCTGTGATTAACAAGATGGCTTTGAAC 60
DB 9 GAACATTCAGATACCTATCATTAATGATGCTGTGATTAACAAGATGGCTTTGAAC 68
QY 61 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
DB 69 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
QY 121 CCTATTCGCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGCAATCCGCTCAGTAC 180
DB 129 CCTATTCGCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGCAATCCGCTCAGTAC 188

QY 181 TACCGTCCCCCGTCCCGAGTACGCGCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGTCCCCCGTCCCGAGTACGCGCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGAGCCCAATCCCATCC 267
 DB 249 GTCTGACGAGCCCAATCCCATCC 275

RESULT 4

US-09-759-143-896
 ; Sequence 896, Application US/09759143
 ; Patent No. US2002002248A1
 ; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, Devin C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yuqi
 APPLICANT: Henderson, Robert A.
 APPLICANT: Kalos, Michael D.
 APPLICANT: Fanger, Gary R.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stolk, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedrick, Thomas S.
 APPLICANT: Carter, Darick
 APPLICANT: Li, Samuel
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.427C23
 CURRENT APPLICATION NUMBER: US/09/759,143
 CURRENT FILING DATE: 2001-01-12
 NUMBER OF SEQ ID NOS: 934
 SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 896
 ; LENGTH: 683
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-759-143-896

Query Match 100.0%; Score 267; DB 10; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5,7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACAGCAAGTGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACAGCAAGTGGCTTTGAAC 68
 QY 61 TCAGGGTACACCAAGCTATTGAGACTTACTATGAAAAACAATGATACCAACCGGAAAAAC 120
 DB 69 TCAGGGTACACCAAGCTATTGAGACTTACTATGAAAAACAATGATACCAACCGGAAAAAC 128
 QY 121 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 180
 DB 129 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 188
 QY 181 TACCGTCCCCCGTCCCGAGTACGCGCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGTCCCCCGTCCCGAGTACGCGCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGAGCCCAATCCCATCC 267
 DB 249 GTCTGACGAGCCCAATCCCATCC 275

RESULT 5

US-09-780-669-896
 ; Sequence 896, Application US/09780669
 ; Patent No. US2002005197A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, Devin C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yuqi
 APPLICANT: Henderson, Robert A.
 APPLICANT: Kalos, Michael D.
 APPLICANT: Fanger, Gary R.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stolk, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedrick, Thomas S.
 APPLICANT: Carter, Darick
 APPLICANT: Li, Samuel
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William
 APPLICANT: Hurrell, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.427C24
 CURRENT APPLICATION NUMBER: US/09/780,669
 CURRENT FILING DATE: 2001-02-09
 NUMBER OF SEQ ID NOS: 943
 SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 896
 ; LENGTH: 683
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-780-669-896

Query Match 100.0%; Score 267; DB 10; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5,7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACAGCAAGTGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACAGCAAGTGGCTTTGAAC 68
 QY 61 TCAGGGTACACCAAGCTATTGAGACTTACTATGAAAAACAATGATACCAACCGGAAAAAC 120
 DB 69 TCAGGGTACACCAAGCTATTGAGACTTACTATGAAAAACAATGATACCAACCGGAAAAAC 128
 QY 121 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 180
 DB 129 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 188
 QY 181 TACCGTCCCCCGTCCCGAGTACGCGCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGTCCCCCGTCCCGAGTACGCGCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGAGCCCAATCCCATCC 267
 DB 249 GTCTGACGAGCCCAATCCCATCC 275

RESULT 6

US-09-822-827-896
 ; Sequence 896, Application US/09822827
 ; Patent No. US20020081680A1
 ; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.534C1
 CURRENT APPLICATION NUMBER: US/09/822,827
 CURRENT FILING DATE: 2001-03-28
 NUMBER OF SEQ ID NOS: 982
 SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 896

LENGTH: 683
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-822-827-896

Query Match 100.0%; Score 267; DB 10; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5,7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATCATCTCCGATGCTGTTGATACAGCAAGATGGCTTGAAC 60
 DB 9 GAACATTCCAGATACCTATCATCTCCGATGCTGTTGATACAGCAAGATGGCTTGAAC 68
 QY 61 TCAGGATCACCACGACCTATTGAGCCTTACTATGAAGAACATGATACCAACCGGAAAC 120
 DB 69 TCAGGATCACCACGACCTATTGAGCCTTACTATGAAGAACATGATACCAACCGGAAAC 128
 QY 121 CCTATCCCGACACGACCTGCTGTTCCCACTGTCTAGAGGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCCGACACGACCTGCTGTTCCCACTGTCTAGAGGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCCCGACGATACGCGAGGCTCTAGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGGTCCCGGTCCCGACGATACGCGAGGCTCTAGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGCACGACGCCAATATCCCATCC 267
 DB 249 GTCTGCACGACGCCAATATCCCATCC 275

RESULT 7

US-10-012-896-894
 Sequence 894, Application US/10012896
 Publication No. US20020183251A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, David C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yugu
 APPLICANT: Kalos, Michael D.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stoik, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Derrick
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aljun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William T.
 APPLICANT: Henderson, Robert A.
 APPLICANT: Hural, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Vinals de Basols, Carlota
 APPLICANT: Foy, Teresa
 APPLICANT: Fanger, Gary R.
 APPLICANT: Mantabane, Yoshihiro
 APPLICANT: Mesher, Nadeline Joy
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.427C27
 CURRENT APPLICATION NUMBER: US/10/012,896
 CURRENT FILING DATE: 2001-12-10
 NUMBER OF SEQ ID NOS: 1011
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 894
 LENGTH: 2479
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-012-896-894

Query Match 100.0%; Score 267; DB 9; Length 2479;

Best Local Similarity 100.0%; Pred. No. 9,4e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATCATCTCCGATGCTGTTGATACAGCAAGATGGCTTGAAC 60
 DB 9 GAACATTCCAGATACCTATCATCTCCGATGCTGTTGATACAGCAAGATGGCTTGAAC 68
 QY 61 TCAGGATCACCACGACCTATTGAGCCTTACTATGAAGAACATGATACCAACCGGAAAC 120
 DB 69 TCAGGATCACCACGACCTATTGAGCCTTACTATGAAGAACATGATACCAACCGGAAAC 128
 QY 121 CCTATCCCGACACGACCTGCTGTTCCCACTGTCTAGAGGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCCGACACGACCTGCTGTTCCCACTGTCTAGAGGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCCCGACGATACGCGAGGCTCTAGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGGTCCCGGTCCCGACGATACGCGAGGCTCTAGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGCACGACGCCAATATCCCATCC 267
 DB 249 GTCTGCACGACGCCAATATCCCATCC 275

RESULT 8

US-09-895-793-894
 Sequence 894, Application US/09895793
 Publication No. US20020192763A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, David C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yugu
 APPLICANT: Kalos, Michael D.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stoik, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Derrick
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aljun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William T.
 APPLICANT: Henderson, Robert A.
 APPLICANT: Hural, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Vinals de Basols, Carlota
 APPLICANT: Foy, Teresa
 APPLICANT: Fanger, Gary R.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.534C2
 CURRENT APPLICATION NUMBER: US/09/895,793
 CURRENT FILING DATE: 2001-06-29
 NUMBER OF SEQ ID NOS: 982
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 894
 LENGTH: 2479
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-895-793-894

Query Match 100.0%; Score 267; DB 9; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 9,4e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATCATCTCCGATGCTGTTGATACAGCAAGATGGCTTGAAC 60
 DB 9 GAACATTCCAGATACCTATCATCTCCGATGCTGTTGATACAGCAAGATGGCTTGAAC 68
 QY 61 TCAGGATCACCACGACCTATTGAGCCTTACTATGAAGAACATGATACCAACCGGAAAC 120

Db 69 TCAGGTACACCAAGTATGACCTTACTATGAAACCAATGATACCAACCGGAAAC 128
Oy 121 CCTATCCCGACACCCCACTGTGTCCCATGTCTAGAGAGTGTATCCGGCTCAATAC 180
Db 129 CCTATCCCGACACCCCACTGTGTCCCATGTCTAGAGAGTGTATCCGGCTCAATAC 188
Oy 181 TACCGTCCCGGTGCGCCGATAGAGCCCGAGAGGTCTGAGACGAGGCTTCAACCCCGTC 240
Db 189 TACCGTCCCGGTGCGCCGATAGAGCCCGAGAGGTCTGAGACGAGGCTTCAACCCCGTC 248
Oy 241 GTCTGACGACGAGCCCAATCCCATCC 267
Db 249 GTCTGACGACGAGCCCAATCCCATCC 275

RESULT 9
US-09-895-814-894
Sequence 894, Application US/09895814
Publication No. US20020193296A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jlang, Yugui
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darlick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C26
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 894
LENGTH: 2479
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-814-894

Query Match 100.0%; Score 267; DB 9; Length 2479;
Best Local Similarity 100.0%; Pred. No. 9.4e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAACATTCGATACCTATGATCTGATGCTGTGATTAACACCAAGATGGCTTTGAAC 60
Db 9 GAACATTCGATACCTATGATCTGATGCTGTGATTAACACCAAGATGGCTTTGAAC 68
Oy 61 TCAGGTACACCAAGTATGACCTTACTATGAAACCAATGATACCAACCGGAAAC 120
Db 69 TCAGGTACACCAAGTATGACCTTACTATGAAACCAATGATACCAACCGGAAAC 128
Oy 121 CCTATCCCGACACCCCACTGTGTCCCATGTCTAGAGAGTGTATCCGGCTCAATAC 180
Db 129 CCTATCCCGACACCCCACTGTGTCCCATGTCTAGAGAGTGTATCCGGCTCAATAC 188
Oy 181 TACCGTCCCGGTGCGCCGATAGAGCCCGAGAGGTCTGAGACGAGGCTTCAACCCCGTC 240
Db 249 GTCTGACGACGAGCCCAATCCCATCC 275

Db 189 TACCGTCCCGGTGCGCCGATAGAGCCCGAGAGGTCTGAGACGAGGCTTCAACCCCGTC 248
Oy 241 GTCTGACGACGAGCCCAATCCCATCC 267
Db 249 GTCTGACGACGAGCCCAATCCCATCC 275

RESULT 10
US-09-759-143-894
Sequence 894, Application US/09759143
Patent No. US20020022248A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jlang, Yugui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darlick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 894
LENGTH: 2479
TYPE: DNA
ORGANISM: Homo sapiens
US-09-759-143-894

Query Match 100.0%; Score 267; DB 10; Length 2479;
Best Local Similarity 100.0%; Pred. No. 9.4e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAACATTCGATACCTATGATCTGATGCTGTGATTAACACCAAGATGGCTTTGAAC 60
Db 9 GAACATTCGATACCTATGATCTGATGCTGTGATTAACACCAAGATGGCTTTGAAC 68
Oy 61 TCAGGTACACCAAGTATGACCTTACTATGAAACCAATGATACCAACCGGAAAC 120
Db 69 TCAGGTACACCAAGTATGACCTTACTATGAAACCAATGATACCAACCGGAAAC 128
Oy 121 CCTATCCCGACACCCCACTGTGTCCCATGTCTAGAGAGTGTATCCGGCTCAATAC 180
Db 129 CCTATCCCGACACCCCACTGTGTCCCATGTCTAGAGAGTGTATCCGGCTCAATAC 188
Oy 181 TACCGTCCCGGTGCGCCGATAGAGCCCGAGAGGTCTGAGACGAGGCTTCAACCCCGTC 240
Db 189 TACCGTCCCGGTGCGCCGATAGAGCCCGAGAGGTCTGAGACGAGGCTTCAACCCCGTC 248
Oy 241 GTCTGACGACGAGCCCAATCCCATCC 267
Db 249 GTCTGACGACGAGCCCAATCCCATCC 275

RESULT 11
US-09-780-669-894
Sequence 894, Application US/09780669
Patent No. US20020051977A1
GENERAL INFORMATION:

Query Match	100.0%	Pred. 267;	DB 10;	Length 2479;
Best Local Similarity	100.0%	Pred. No. 9,4e-75;		
Matches 267; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 12
US-09-822-827-894

```

1 Sequence 894, Application US/09822827
2 Patent No. US20020081680A1
3 GENERAL INFORMATION:
4 APPLICANT: Xu, Jiangchun
5 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
6 TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
7 FILE REFERENCE: 210121.534C1
8 CURRENT APPLICATION NUMBER: US/09/822,827
9 CURRENT FILING DATE: 2001-03-28
10 NUMBER OF SEQ ID NOS: 982
11 SOFTWARE: FASTSEQ for Windows Version 3.0
12 SEQ ID NO 894
13 LENGTH: 2479

```

Query Match	100.0%	Score 267;	DB 10;	Length 2479;
Best Local Similarity	100.0%	Pred. No. 9.4e-75;		
Matches 267;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

RESULT 13
US-10-012-896-925

; Sequence 929, Application US/100128966
; Publication No. US20020183251A1

```

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bascois, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Wantanabe, Yoshihiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 20121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 929
LENGTH: 3245
TYPE: DNA
ORGANISM: Homo sapiens
US-10-012-896-929

```

Query Match	100.0%;	Score 267;	DB 9;	Length 3245;
Best Local Similarity	100.0%;	Pred. No. 1e-74;		

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCGAGATACCTATCTATCTGATGCTTTGATTAACAGACAGATGCTTTGAAC 60
DB 97 GAACATTCGAGATACCTATCTATCTGATGCTTTGATTAACAGACAGATGCTTTGAAC 156

QY 61 TCAGGGTCAACACAGCTATTTGACCTTATCTATGATAAACCATTGATACCAACCGGAAAC 120
DB 157 TCAGGGTCAACACAGCTATTTGACCTTATCTATGATAAACCATTGATACCAACCGGAAAC 216

QY 121 CCTATATCCCGACAGCCCACTGTGCTCCCACTGTCTAAGAGTGATCCGGCTCAGTAC 180
DB 217 CCTATATCCCGACAGCCCACTGTGCTCCCACTGTCTAAGAGTGATCCGGCTCAGTAC 276

QY 181 TACCCGTCCTCCCGTCCCGCAGTACGCCCCGAGGGTCTTGAAGCAGGCTTCCAAACCCCGTC 240
DB 277 TACCCGTCCTCCCGTCCCGCAGTACGCCCCGAGGGTCTTGAAGCAGGCTTCCAAACCCCGTC 336

QY 241 GTCTGACAGCAGCCCAATCCCATCC 267
DB 337 GTCTGACAGCAGCCCAATCCCATCC 363

RESULT 14

US-09-895-793-929
Sequence 929, Application US/09895793
Publication No. US20020192763A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.534C2
CURRENT APPLICATION NUMBER: US/09/895,793
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 929
LENGTH: 3245
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-793-929

Query Match 100.0%; Score 267; DB 9; Length 3245;
Best Local Similarity 100.0%; Pred. No. 1e-74; Indels 0; Gaps 0;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCGAGATACCTATCTATCTGATGCTTTGATTAACAGACAGATGCTTTGAAC 60
DB 97 GAACATTCGAGATACCTATCTATCTGATGCTTTGATTAACAGACAGATGCTTTGAAC 156

QY 61 TCAGGGTCAACACAGCTATTTGACCTTATCTATGATAAACCATTGATACCAACCGGAAAC 120
DB 157 TCAGGGTCAACACAGCTATTTGACCTTATCTATGATAAACCATTGATACCAACCGGAAAC 216

DB 157 TCAGGGTCAACACAGCTATTTGACCTTATCTATGATAAACCATTGATACCAACCGGAAAC 216
QY 121 CCTATATCCCGACAGCCCACTGTGCTCCCACTGTCTAAGAGTGATCCGGCTCAGTAC 180
DB 217 CCTATATCCCGACAGCCCACTGTGCTCCCACTGTCTAAGAGTGATCCGGCTCAGTAC 276

QY 181 TACCCGTCCTCCCGTCCCGCAGTACGCCCCGAGGGTCTTGAAGCAGGCTTCCAAACCCCGTC 240
DB 277 TACCCGTCCTCCCGTCCCGCAGTACGCCCCGAGGGTCTTGAAGCAGGCTTCCAAACCCCGTC 336

QY 241 GTCTGACAGCAGCCCAATCCCATCC 267
DB 337 GTCTGACAGCAGCCCAATCCCATCC 363

RESULT 15

US-09-895-814-929
Sequence 929, Application US/09895814
Publication No. US20020193296A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C26
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 929
LENGTH: 3245
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-814-929

Query Match 100.0%; Score 267; DB 9; Length 3245;
Best Local Similarity 100.0%; Pred. No. 1e-74; Indels 0; Gaps 0;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCGAGATACCTATCTATCTGATGCTTTGATTAACAGACAGATGCTTTGAAC 60
DB 97 GAACATTCGAGATACCTATCTATCTGATGCTTTGATTAACAGACAGATGCTTTGAAC 156

QY 61 TCAGGGTCAACACAGCTATTTGACCTTATCTATGATAAACCATTGATACCAACCGGAAAC 120
DB 157 TCAGGGTCAACACAGCTATTTGACCTTATCTATGATAAACCATTGATACCAACCGGAAAC 216

QY 121 CCTATATCCCGACAGCCCACTGTGCTCCCACTGTCTAAGAGTGATCCGGCTCAGTAC 180
DB 217 CCTATATCCCGACAGCCCACTGTGCTCCCACTGTCTAAGAGTGATCCGGCTCAGTAC 276

QY 181 TACCCGTCCTCCCGTCCCGCAGTACGCCCCGAGGGTCTTGAAGCAGGCTTCCAAACCCCGTC 240
DB 277 TACCCGTCCTCCCGTCCCGCAGTACGCCCCGAGGGTCTTGAAGCAGGCTTCCAAACCCCGTC 336

Db 277 TACCGTCCCCCGTGGCCCACTACGCCCCGAGGGTCTGACGACAGGCTTCCACCCTGTC 336
Qy 241 GTCTGACGACAGCCCAATCCCATCC 267
Db 337 GTCTGACGACAGCCCAATCCCATCC 363

Search completed: January 31, 2003, 09:39:12
Job time : 76 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 06:47:46 ; Search time 308 Seconds
(without alignments)
1952.219 Million cell updates/sec

Title: US-09-807-201-7
Perfect score: 267
Sequence: 1 Gacattccagatctatc.....cgccagcccaatccatcc 267

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq.101002:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	267	100.0	267 21 AA295004	CSG Prolis EST use
2	267	100.0	683 22 AAS64165	Human CDNA encodin
3	267	100.0	683 22 AAH93929	Human transmembran
4	267	100.0	683 24 ABL95536	Human transmembran
5	267	100.0	1738 21 AA229636	Human 20PFI2-GTC2
6	267	100.0	1738 24 AAD28778	Human 20PFI2-GTC1
7	267	100.0	2479 21 AA290478	Ovrlis homolog pro
8	267	100.0	2479 21 AA287813	Human tumour suppr
9	267	100.0	2479 22 AAS64164	Human CDNA encodin

10	267	100.0	2479 22 AAD13168	Human serine prote
11	267	100.0	2479 22 AAH93928	Human transmembran
12	267	100.0	2479 24 ABR93201	Prostate cancer-as
13	267	100.0	2479 24 ABL95535	Human transmembran
14	267	100.0	2479 24 AAD28779	Human TMPS52 gene
15	267	100.0	3245 21 AAA08803	Androgen-inducible
16	267	100.0	3245 22 AAS64178	Human prostate CDN
17	267	100.0	3245 22 AAH93942	Human prostate CDN
18	267	100.0	3245 24 ABL95549	Human P100C full length
19	267	100.0	3483 21 AA295005	Human P100C CDNA
20	267	100.0	3483 23 ABV21748	Cancer specific ge
21	267	100.0	3483 23 ABV21767	Human prostate exp
22	267	100.0	3483 23 ABV2318	Human prostate exp
23	267	100.0	3483 23 ABV24651	Human prostate exp
24	267	100.0	3483 23 ABV27570	Human prostate exp
25	267	100.0	3483 23 ABV27589	Human prostate exp
26	267	100.0	3483 23 ABV29119	Human prostate exp
27	267	100.0	3483 23 ABV29165	Human prostate exp
28	267	100.0	3483 23 ABL95515	Human prostate exp
29	267	100.0	3483 21 AAC83325	Human TMPS52 DNA
30	267	100.0	3483 22 AAS64180	Human TMPS52 DNA
31	267	100.0	3483 22 AAH93944	Human prostate CDN
32	267	100.0	3483 24 ABL95551	Human P100C open readin
33	267	100.0	3483 21 AA287786	Human P100C CDNA
34	267	100.0	3483 22 AAS64179	Human tumour suppr
35	267	100.0	3483 22 AAH93943	Human prostate CDN
36	267	100.0	3483 24 ABL95550	Human P100C open readin
37	267	100.0	3483 21 AAC06645	Human P100C CDNA
38	267	100.0	3483 19 AAV61239	Human secreted pro
39	267	100.0	3483 19 AAV56624	CDNA sequence of p
40	267	100.0	3483 21 AAA06587	Prostate tumour sp
41	267	100.0	3483 22 AAS63595	Human immunogenic
42	267	100.0	3483 22 AAS10146	Human prostate CDN
43	267	100.0	3483 22 AAH93503	Human prostate tum
44	267	100.0	3483 22 AAH84817	Human prostate-spe
45	267	100.0	3483 22 AAH02568	Human prostate-spe
			196 24 ABL94967	Prostate tumour an
				Human P255 CDNA se

ALIGNMENTS

RESULT 1
ID AA295004 standard; CDNA; 267 BP.
XX
AC AA295004;
XX
DT 15-AUG-2000 (first entry)
XX
DE CSG Prolis EST useful as prostate cancer marker.
XX
KW Prostate cancer; cancer specific gene; CSG; expressed sequence tag;
KW EST; diagnosis; monitoring; staging; imaging; therapy; metastasis;
KW marker; human; Prolis; ss.
XX
OS Homo sapiens.
XX
PN WO200023111-A1.
XX
PD 27-APR-2000.
XX
PF 19-OCT-1999; 99WO-US24331.
XX
PR 19-OCT-1998; 98US-0104737.
XX
FA (DIAD-) DIADEXUS LLC.
XX
PI Salceda S, Recipon H, Catterkey R;
XX
DR WPI; 2000-339531/29.
XX
PT Diagnosing, staging and monitoring the presence and metastases of
PT prostate cancer especially useful for treating prostate cancer

PT comprises measuring changes in cancer specific gene levels -

XX Claim 7; Page 57; 74pp; English.

PS

XX

CC The present sequence is that of cancer specific gene (CSG) expressed

CC sequence tag clone 2189835H1 (Prolis); the full-length contig is

CC given in AA295005. The CSG was identified in a database search using

CC the data mining Cancer Leads Automatic Search Package (CLASP), which

CC allows the identification of highly expressed organ and cancer

CC specific genes. Overexpression of Prolis was observed in 3 of 4

CC primary prostate cancer tissues examined, indicative of it being a

CC diagnostic marker for prostate cancer. The invention provides BSA

CC and full-length contigs for prostate CSGs (see AA29498-295017). The

CC CSGs, polypeptides encoded by them, and antibodies that specifically

CC bind CSG are used in new, claimed methods for detecting, diagnosing,

CC monitoring, staging, imaging and treating prostate cancer. The new

CC methods provide earlier diagnosis for the presence and metastasis

CC of prostate cancer, and can be used to determine if a cancer has

CC metastasized, or to monitor the progress or stage of the disease

CC when it has not metastasized.

XX

SQ Sequence 267 BP; 64 A; 97 C; 53 G; 53 T; 0 other;

Query Match 100.0%; Score 267; DB 21; Length 267;

Best Local Similarity 100.0%; Pred. No. 2.4e-73;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACATCATATCTAGTGGTGTGATAGCAAGATGGCTTGAAC 60

DB 1 GAACATTCAGATACATCATATCTAGTGGTGTGATAGCAAGATGGCTTGAAC 60

QY 61 TCAGGGTCAACACAGCTATTGGACCTTACTATGAAAACCATGATACCAACCGAANAAC 120

DB 61 TCAGGGTCAACACAGCTATTGGACCTTACTATGAAAACCATGATACCAACCGAANAAC 120

QY 121 CCTATATCCGCAAGCCCACTGTGTCCCACTGTCTTACAGAGTGATCGGCTCAGTAC 180

DB 121 CCTATATCCGCAAGCCCACTGTGTCCCACTGTCTTACAGAGTGATCGGCTCAGTAC 180

QY 181 TACCCGTCCCGCGTCCGAGTACGCGCCGAGGGTCTTACGAGAGCTTCCAAACCCCGTC 240

DB 181 TACCCGTCCCGCGTCCGAGTACGCGCCGAGGGTCTTACGAGAGCTTCCAAACCCCGTC 240

QY 241 GTCTGCAGCGACGCCAATATCCCATCC 267

DB 241 GTCTGCAGCGACGCCAATATCCCATCC 267

RESULT 2

AA564165

ID AA564165 standard; cDNA; 683 BP.

XX

AC AA564165;

XX

DT 29-JAN-2002 (first entry)

XX

DE Human cDNA encoding transmembrane serine protease 2 amino acids 1-209.

XX

KM Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.

OS Homo sapiens.

XX

PN WO200173032-A2.

XX

PD 04-OCT-2001.

XX

PF 27-MAR-2001; 2001WO-US09919.

XX

PR 27-MAR-2000; 2000US-0536857.

XX

PR 09-MAY-2000; 2000US-0568100.

XX

PR 12-MAY-2000; 2000US-0570737.

XX

PR 13-JUN-2000; 2000US-0593793.

XX

PR 27-JUN-2000; 2000US-0605783.

PR 10-AUG-2000; 2000US-0636215.

PR 29-AUG-2000; 2000US-0651236.

PR 06-SEP-2000; 2000US-0657279.

PR 02-OCT-2000; 2000US-0679426.

PR 10-OCT-2000; 2000US-0685166.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Ketter MW, Stoik JA, Day CH, Vedvick TS, Carter D;

PI Li SX, Wang A, Skelky YAW, Hepler WT, Henderson RA;

DR WPI; 2001-639232/73.

PS P-PSDB; AAU69940.

XX

PT New human prostate-specific polypeptides and polynucleotides useful for

PT the diagnosis and treatment of cancer, especially prostate cancer -

XX

PS Claim 1; Page 557; 579pp; English.

XX

CC The invention relates to isolated prostate-specific

CC polynucleotides, polypeptides, fusion proteins of the polypeptides,

CC antibodies raised against the polypeptides (or antigenic epitopes

CC derived from them) and antigen-presenting cells expressing the

CC polypeptides. The antibodies are useful for detecting the presence of

CC cancer, especially prostate cancer. The polypeptides, polynucleotides and

CC the antigen-presenting cells are useful for stimulating and/or expanding

CC T cells specific for a tumour protein, and for inhibiting the development

CC of cancer especially prostate cancer. Compositions comprising the

CC polynucleotide and/or polypeptide are useful for stimulating an immune

CC response, and for treating cancer. The oligonucleotide is useful for

CC detecting cancer. The present sequence is a prostate specific

CC polynucleotide of the invention.

XX

SQ Sequence 683 BP; 160 A; 209 C; 171 G; 143 T; 0 other;

Query Match 100.0%; Score 267; DB 22; Length 683;

Best Local Similarity 100.0%; Pred. No. 3.3e-73;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACATCATATCTAGTGGTGTGATAGCAAGATGGCTTGAAC 60

DB 9 GAACATTCAGATACATCATATCTAGTGGTGTGATAGCAAGATGGCTTGAAC 68

QY 61 TCAGGGTCAACACAGCTATTGGACCTTACTATGAAAACCATGATACCAACCGAANAAC 120

DB 69 TCAGGGTCAACACAGCTATTGGACCTTACTATGAAAACCATGATACCAACCGAANAAC 128

QY 121 CCTATATCCGCAAGCCCACTGTGTCCCACTGTCTTACAGAGTGATCGGCTCAGTAC 180

DB 129 CCTATATCCGCAAGCCCACTGTGTCCCACTGTCTTACAGAGTGATCGGCTCAGTAC 188

QY 181 TACCCGTCCCGCGTCCGAGTACGCGCCGAGGGTCTTACGAGAGCTTCCAAACCCCGTC 240

DB 189 TACCCGTCCCGCGTCCGAGTACGCGCCGAGGGTCTTACGAGAGCTTCCAAACCCCGTC 248

QY 241 GTCTGCAGCGACGCCAATATCCCATCC 267

DB 249 GTCTGCAGCGACGCCAATATCCCATCC 275

RESULT 3

AAH93929

ID AAH93929 standard; cDNA; 683 BP.

XX

AC AAH93929;

XX

DT 04-OCT-2001 (first entry)

XX

DE Human transmembrane protease serine 2 partial cDNA sequence.

XX

KM Human; prostate cancer; prostate-specific; diagnosis; vaccine;

KM cytostatic; gene therapy; metastasis; ss.

XX OS Homo sapiens.
 XX PN MO200151633-A2.
 XX PD 19-JUL-2001.
 XX PF 16-JAN-2001; 2001MO-US01574.
 XX PR 14-JAN-2000; 2000US-0483572.
 XX PA (CORI-) CORIXA CORP.
 XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 XX PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW,
 XX PI Wang A, Meagher MJ;
 XX XX WPI; 2001-425873/45.
 XX PT New polynucleotide encoding a prostate-specific protein, for
 XX PT diagnosing, monitoring and treating prostate cancer in a patient and
 XX PT for use in vaccines -
 XX PS Claim 1; Page 521; 543pp; English.
 XX XX
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II) (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93157 to AAH93944 and AA01115 to
 CC AA01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.
 XX SO Sequence 683 BP, 160 A, 209 C, 171 G, 143 T, 0 other;
 Query Match 100.0%; Score 267; DB 22; Length 683;
 Best Local Similarity 100.0%; Pred. No. 3.3e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGCTTGAAC 60
 DB 9 GAACATTCGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGCTTGAAC 68
 QY 61 TCAAGGTACACCAAGCTATTGGACCTTACTATGAAAAACCATGATACCAACCGAAAAAC 120
 DB 69 TCAAGGTACACCAAGCTATTGGACCTTACTATGAAAAACCATGATACCAACCGAAAAAC 128
 QY 121 CCTATCCGCGACAGCGCATGTGGTCCCATGCTAGAGGTGATCCGGTCACTAC 180
 DB 129 CCTATCCGCGACAGCGCATGTGGTCCCATGCTAGAGGTGATCCGGTCACTAC 188
 QY 181 TACCGGTCCCCCGTCCCGAGTACGCCCCGAGAGGTCCTGACGCAAGGCTTCAACCCCGTC 240
 DB 189 TACCGGTCCCCCGTCCCGAGTACGCCCCGAGAGGTCCTGACGCAAGGCTTCAACCCCGTC 248
 QY 241 GTCTGACGACGACCCCAATCCCATCC 267
 DB 249 GTCTGACGACGACCCCAATCCCATCC 275
 RESULT 4
 ABL95536
 ID ABL95536 standard; cDNA, 683 BP.
 XX AC ABL95536;
 XX

DT 19-JUL-2002 (first entry)
 XX XX
 DE Human transmembrane protease serine 2 cDNA fragment SEQ ID NO 896.
 XX XX
 KM Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
 XX XX gene therapy; gene; ss.
 XX OS Homo sapiens.
 XX PN US2002022248-A1.
 XX PD 21-FEB-2002.
 XX XX
 PF 12-JAN-2001; 2001US-0759143.
 XX XX
 XX 25-FEB-1997; 97US-0806099.
 PR 01-AUG-1997; 97US-0904804.
 PR 09-FEB-1998; 98US-0020956.
 PR 25-FEB-1998; 98US-0030607.
 PR 14-JUL-1998; 98US-0115453.
 PR 23-SEP-1998; 98US-0159812.
 PR 15-JAN-1999; 99US-0232149.
 PR 09-APR-1999; 99US-0288946.
 PR 13-JUL-1999; 99US-0352616.
 PR 12-NOV-1999; 99US-0439313.
 PR 18-NOV-1999; 99US-0443686.
 PR 14-JAN-2000; 2000US-0483672.
 PR 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX XX
 PA (XUJ/) XU J.
 PA (DILL/) DILLON D C.
 PA (MITC/) MITCHAM J L.
 PA (HARL/) HARLOCKER S L.
 PA (JIANG/) JIANG Y.
 PA (KALOS/) KALOS M D.
 PA (FANG/) FANGER G R.
 PA (RETT/) RETTER M W.
 PA (STOL/) STOLK J A.
 PA (DAYC/) DAY C H.
 PA (VEDV/) VEDVICK T S.
 PA (CART/) CARTER D.
 PA (LISK/) LI S X.
 PA (WANG/) WANG A.
 PA (SKEI/) SKEIKY Y A W.
 PA (HEPL/) HEPLER W T.
 PA (HEND/) HENDERSON R A.
 XX XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX XX WPI; 2002-255649/30.
 XX DR
 XX PT New prostate-specific polynucleotides for diagnosing and treating
 XX PT diseases, in particular prostate cancer, and as markers for the
 XX PT progression of cancer -
 XX XX
 PS Claim 1; SEQ ID NO 896; 87pp; English.
 XX XX
 CC The present invention provides prostate-specific coding sequences and
 CC their encoded proteins. These can be used in the diagnosis and treatment
 CC of cancer, particularly prostate cancer. The present sequence is a cDNA
 CC described in the invention.
 XX XX

Sequence 683 BP; 160 A; 209 C; 171 G; 143 T; 0 other;

Query Match 100.0%; Score 267; DB 24; Length 683;
Best Local Similarity 100.0%; Pred. No. 3.3e-73;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 60
DB 9 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 68
QY 61 TCAGGGTCAACGACGACTATGATGATGAAACCATGATACCAACCGGAAAC 120
DB 69 TCAGGGTCAACGACGACTATGATGATGAAACCATGATACCAACCGGAAAC 128
QY 121 CCTATCCCGACAGCCCACTGTGTCCTGCTATGAGAGTCATCCGGCTCAGTAC 180
DB 129 CCTATCCCGACAGCCCACTGTGTCCTGCTATGAGAGTCATCCGGCTCAGTAC 188
QY 181 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTGACCGAGCTTCAACCCCGTC 240
DB 189 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTGACCGAGCTTCAACCCCGTC 248
QY 241 GTCTGACGACGACCAATCCCATCC 267
DB 249 GTCTGACGACGACCAATCCCATCC 275

RESULT 5
AA229636
ID AA229636 standard; cDNA; 1738 BP.

AC AA229636;
DT 22-MAR-2000 (first entry)

DE Human 20P1F12-GTCC2 encoding cDNA.

KW 20P1F12; TMPRSS2; androgen; serine protease; 20P1F12-GTCC1; cancer;
KM transmembrane protein; colon; prostate; prostate tumour; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 112..1590
FT CDS /tag= a
FT /product= "20P1F12-GTCC2"
FT /note= "Prostate-specific serine protease"

PN MO9962942-A2.

XX 09-DEC-1999.

PF 01-JUN-1999; 99WO-US12253.

PR 01-JUN-1998; 98US-0087598.

PR 29-JUN-1998; 98US-0091474.

PR 14-APR-1999; 99US-0129521.

XX (UROC-) UROGENESYS INC.

PA (APAR) APAR D E.

PA (HUBE) HUBERT R S.

PA (LEON) LEONG K.

PA (RAIT) RAITANO A B.

PA (SAFE) SAFRAN D C.

PI Afar DE, Hubert RS, Leong K, Raitano AB, Safiran DC;

DR WPI, 2000-116363/10.

DR P-PDB; AAY44406.

Novel cell surface antigen useful to treat colon and prostate cancer -
Claim 2, Fig 1, 58pp, English.

XX The present sequence is a cDNA encoding 20P1F12 protein (also known as
CC the TMPRSS2 protein), which is a prostate-specific, androgen-regulated,
CC cell surface serine protease. It is a glycosylated type II transmembrane
CC protein with an extracellular C-terminal protease domain. Host cells can
CC be transformed to produce the protein, using vector containing
CC 20P1F12/TMPRSS2 gene (also designated 20P1F12-GTCC), as deposited with
CC ATCC accession number 207097). Anti-20P1F12/TMPRSS2 antibodies may be
CC used as therapeutic agent for prostate and colon cancers, to image
CC prostate cancer cells and prostate tumours, to identify ligands and
CC cellular constituents that bind to a 20P1F12/TMPRSS2 gene product and for
CC use as cancer vaccines.

Sequence 1738 BP; 433 A; 459 C; 473 G; 373 T; 0 other;

Query Match 100.0%; Score 267; DB 21; Length 1738;
Best Local Similarity 100.0%; Pred. No. 4.6e-73;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 60
DB 64 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 123
QY 61 TCAGGGTCAACGACGACTATGATGATGAAACCATGATACCAACCGGAAAC 120
DB 124 TCAGGGTCAACGACGACTATGATGATGAAACCATGATACCAACCGGAAAC 183
QY 121 CCTATCCCGACAGCCCACTGTGTCCTGCTATGAGAGTCATCCGGCTCAGTAC 180
DB 184 CCTATCCCGACAGCCCACTGTGTCCTGCTATGAGAGTCATCCGGCTCAGTAC 243
QY 181 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTGACCGAGCTTCAACCCCGTC 240
DB 244 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTGACCGAGCTTCAACCCCGTC 303
QY 241 GTCTGACGACGACCAATCCCATCC 267
DB 304 GTCTGACGACGACCAATCCCATCC 330

RESULT 6
AAD28778
ID AAD28778 standard; cDNA; 1738 BP.

XX AAD28778;

DT 07-MAY-2002 (first entry)

DE Human 20P1F12-GTCC1 cDNA.

KW Serine protease; 20P1F12/TMPRSS2; 20P1F12-GTCC1; cell growth; neoplasm;
KM cancer; vaccine; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 112..1591
FT CDS /tag= a
FT /product= "Human 20P1F12-GTCC1 protein"

PN WO200204953-A2.

XX 17-JAN-2002.

PF 12-JUL-2001; 2001WO-US22168.

PR 12-JUL-2000; 2000US-0615285.

PA (AGEN-) AGENSYS INC.

PI Safran D, Raitano AB, Hubert RS, Jakobovits A, Faris M;

PI Chalfante-Bid PM;

DR WPI, 2002-154967/20.
DR P-PSDB; AA218096.

XX Examining a biological sample for evidence of dysregulated cellular
PT growth, comprises comparing the status of prostate-specific,
XX androgen-regulated, secreted serine protease, 20P1F12/TMPRSS2, in a
PT corresponding normal sample

XX Example 3, Fig 1, 161pp; English.

XX The present invention relates to methods and compositions for the
CC diagnosis and therapy of prostate, colon, bladder, lung, ovarian and
CC kidney cancer derived from or based on a normally prostate-specific,
CC androgen regulated, cell membrane associated secreted serine protease
CC termed 20P1F12/TMPRSS2. The invention further relates to a method of
CC examining a biological sample for evidence of dysregulated cellular
CC growth comprising comparing the status of 20P1F12/TMPRSS2 gene (also
CC designated 20P1F12-GTCL) in the sample to the status of 20P1F12/TMPRSS2
CC in a corresponding normal sample. The invention also relates to 20P1F12/
CC TMPRSS2 polynucleotides and their corresponding proteins. Methods of the
CC invention are used for examining a sample such as blood, serum, stool,
CC urine, semen, or biopsy tissue for evidence of dysregulated cell growth.
CC The dysregulated cell growth is indicative of bladder cancer, lung
CC cancer, kidney cancer or ovarian cancer. It is useful for identifying
CC evidence of a neoplasm in a sample. Vaccines comprising an immunogenic
CC portion of 20P1F12/TMPRSS2 are useful for inhibiting growth of a cell
CC expressing 20P1F12/TMPRSS2 in a patient suffering from bladder cancer,
CC lung cancer, ovarian cancer or metastatic cancer. The present sequence
CC is human 20P1F12-GTCL cDNA.

XX Sequence 1738 BP; 433 A; 459 C; 473 G; 373 T; 0 other;

XX Query Match 100.0%; Score 267; DB 24; Length 1738;

XX Best Local Similarity 100.0%; Pred. No. 4.6e-73; Mismatches 0; Indels 0; Gaps 0;

XX Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATTAACAGCAAGATGGCTTGAAC 60
DB 64 GAACATTCAGATACCTATCTATCTGATGCTGTGATTAACAGCAAGATGGCTTGAAC 123
QY 61 TACGGGTACACCAAGCTATTTGACCTTATGTAACCAATGATACCAACCGGAAAC 120
DB 124 TACGGGTACACCAAGCTATTTGACCTTATGTAACCAATGATACCAACCGGAAAC 183
QY 121 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 180
DB 184 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 243
QY 181 TACCCGTCCCGTGCAGTACGCCCCGAGGGTCTCTACAGAGGTTCCAAACCCCGTC 240
DB 244 TACCCGTCCCGTGCAGTACGCCCCGAGGGTCTCTACAGAGGTTCCAAACCCCGTC 303
QY 241 GTCTGACGACGACCAATTCCTCCATCC 267
DB 304 GTCTGACGACGACCAATTCCTCCATCC 330

RESULT 7
AA290478 standard; cDNA; 2479 BP.

XX AA290478;

XX 06-JUN-2000 (first entry)

XX Ovr115 homolog protein encoding cDNA.

XX CSg; cancer specific gene; cancer; gynecologic cancer; ovarian; breast;
XX endometrial; uterine; lung; cytotoxic; ss.

XX Homo sapiens.
XX MO200012758-A1.

XX 09-MAR-2000.

XX 01-SEP-1999; 99MO-US19655.

XX 02-SEP-1998; 98US-0098880.

XX (DIAD-) DIADEXUS LLC.

XX Salceda S, Sun Y, Recipon H, Caffery R;

XX WPI; 2000-256657/22.

XX P-PSDB; AA57280.

XX Diagnosing, staging, monitoring, imaging and treating cancer especially
PT gynecological cancers e.g. breast, ovarian cancer and lung cancer.
PT Involves measuring cancer specific gene levels in cells and body fluids
PT
PS Claim 9, Page 49-50; 58pp; English.

XX The invention relates to detecting, diagnosing metastasis and staging
CC cancer by measuring levels of cancer specific genes (CSG) in cells,
CC tissues or body fluids. Their remission and progression, decreases and
CC increases in CSG levels, is also monitored, by periodic sample analysis.
CC The methods are useful for detecting cancers, especially gynecologic
CC cancers which include ovarian, breast, endometrial and uterine cancer
CC and lung cancer. Antibodies against the CSGs labeled with paramagnetic
CC ions or a radioisotope is useful for imaging cancer and when conjugated
CC with a cytotoxic agent are useful for treating cancer. The present
CC sequence represents a Ovr115 homolog protein encoding cDNA, that
CC can be used for the detection of the various cancers.

XX Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;

XX Query Match 100.0%; Score 267; DB 21; Length 2479;

XX Best Local Similarity 100.0%; Pred. No. 5.2e-73; Mismatches 0; Indels 0; Gaps 0;

XX Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATTAACAGCAAGATGGCTTGAAC 60
DB 9 GAACATTCAGATACCTATCTATCTGATGCTGTGATTAACAGCAAGATGGCTTGAAC 68
QY 61 TACGGGTACACCAAGCTATTTGACCTTATGTAACCAATGATACCAACCGGAAAC 120
DB 69 TACGGGTACACCAAGCTATTTGACCTTATGTAACCAATGATACCAACCGGAAAC 128
QY 121 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 180
DB 129 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 188
QY 181 TACCCGTCCCGTGCAGTACGCCCCGAGGGTCTCTACAGAGGTTCCAAACCCCGTC 240
DB 189 TACCCGTCCCGTGCAGTACGCCCCGAGGGTCTCTACAGAGGTTCCAAACCCCGTC 248
QY 241 GTCTGACGACGACCAATTCCTCCATCC 267
DB 249 GTCTGACGACGACCAATTCCTCCATCC 275

RESULT 8
AA287813 standard; DNA; 2479 BP.

XX AA287813;

XX 12-MAY-2000 (first entry)

XX Human tumour suppressor TMPRSS2 gene (GenBank Accn No: U75329).

XX Tumour suppressor gene; TMPRSS2; cancer; human; drug design;
XX gene therapy; protein therapy; ds.

OS	Homo sapiens.
XX	
PN	WO200000605-A1.
XX	
PD	06-JAN-2000.
XX	
PF	29-JUN-1999; 99WO-US14622.
XX	
PR	29-JUN-1998; .98US-0091044.
XX	
PA	(MIRI-) MIRIAD GENETICS INC.
XX	
PI	Wong AKC, Tavligian SV, Teng DHF;
XX	
DR	WPI; 2000-170914/15.
XX	
PT	Novel tumor suppressor TMRSS2 used for the diagnosis and prognosis of
PT	human cancer.
XX	
PS	Example 1; Page 84-85; 89pp; English.
XX	
CC	The invention provides a new tumor suppressor gene, designated TMRSS2.
CC	The TMRSS2 polynucleotides and polypeptides can be used in methods for
CC	diagnosing and prognosing predisposition to cancer in humans. The
CC	polypeptides may also be used in assays to screen for compounds with
CC	anti-cancer or therapeutic properties. The polypeptides are also useful
CC	for rational drug design. The TMRSS2 polynucleotides and polypeptides
CC	may be used for gene therapy and protein therapy. The present sequence
CC	represents the TMRSS2 gene (Genbank Accn No: U75329).
SQ	Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
	Query Match 100.0%; Score 267; DB 21; Length 2479;
	Best Local Similarity 100.0%; Pred. No. 5.2e-73;
	Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY	1 GAACATTCGAGTACTCTATCATTTACTCGATGCTGTGTGAIPAACGAGAATGGCTTGAAC 60
DB	9 GAACATTCGAGTAACCTATCATTTACTCGATGCTGTGTGAIPAAACGAGAATGGCTTGAAC 68
QY	61 TAGAGTGCAACACACAGCTATTGGAGCTTGTCTATGAAACCATGAGATAACAACCGAANAAC 120
DB	69 TAGAGTGCAACACACAGCTATTGGAGCTTGTCTATGAAACCATGAGATAACAACCGAANAAC 128
QY	121 CCTATCCGCACAGGCCCACTGTGTGCCCACTGTCTACGAGGTGCATCCGGCTCAGTAC 180
DB	129 CCTATCCGCACAGGCCCACTGTGTGCCCACTGTCTACGAGGTGCATCCGGCTCAGTAC 188
QY	181 TACCGTCCCCCGTGGCCCAAGTACGCCCCGAGGGTCTCTAGCCGAGCTTCCAACCCCGTC 240
DB	189 TACCGTCCCCCGTGGCCCAAGTACGCCCCGAGGGTCTCTAGCCGAGCTTCCAACCCCGTC 248
QY	241 GTCTGACGACGAGCCCAATCCCCTATCC 267
DB	249 GTCTGACGACGAGCCCAATCCCCTATCC 275
RESULT 9	
AAS64164	
ID	AAS64164 standard; cDNA; 2479 BP.
XX	
AC	AAS64164;
XX	
DT	29-JAN-2002 (first entry)
DE	Human cDNA encoding transmembrane serine protease 2.
XX	
KW	Human; prostate cancer; ss; cytosolic; immunostimulant; tumour.
XX	
OS	Homo sapiens.
XX	
PN	WO200173032-A2.

PD	04-OCT-2001.
XX	
PF	27-MAR-2001; 2001WO-US09919.
PR	27-MAR-2000; 2000US-0536857.
XX	
PR	09-MAY-2000; 2000US-0568100.
PR	12-MAY-2000; 2000US-0570737.
PR	13-JUN-2000; 2000US-0593793.
PR	27-JUN-2000; 2000US-0605783.
PR	10-AUG-2000; 2000US-0636213.
PR	29-AUG-2000; 2000US-0651236.
PR	06-SEP-2000; 2000US-0657279.
PR	02-OCT-2000; 2000US-0679426.
PR	10-OCT-2000; 2000US-0685166.
XX	
PA	(CORI-) CORIXA CORP.
XX	
Pt	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
Pt	Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
Pt	Ll SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX	
DR	WPI; 2001-639232/73.
XX	P-FEDB; AAU69939.
PS	New human prostate-specific polypeptides and polynucleotides useful for the diagnosis and treatment of cancer, especially prostate cancer - Claim 1; Page 555-556; 579pp; English.
CC	The invention relates to isolated prostate-specific
CC	polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC	antibodies raised against the polypeptides (or antigenic epitopes
CC	derived from them) and antigen-presenting cells expressing the
CC	polypeptides. The antibodies are useful for detecting the presence of
CC	cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC	the antigen-presenting cells are useful for stimulating and/or expanding
CC	T cells specific for a tumour protein, and for inhibiting the development
CC	of cancer especially prostate cancer. Compositions comprising the
CC	polynucleotide and/or polypeptide are useful for stimulating an immune
CC	response, and for treating cancer. The oligonucleotide is useful for
CC	detecting cancer. The present sequence is a prostate specific
CC	polynucleotide of the invention.
XX	
SQ	Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
	Query Match 100.0%; Score 267; DB 22; Length 2479;
	Best Local Similarity 100.0%; Pred. NO. 5.2e-73;
	Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 GAACATTCAGATCCTTATCATTTACTGATGCTGTGGATAACAGCAAGATGGCTTGAAC 60
DB	9 GAACATTCAGATCCTTATCATTTACTGATGCTGTGGATAACAGCAAGATGGCTTGAAC 68
OY	61 TCAGGTCACCAACCAGCTATTGAACTTAATGTGAAAACATGATGCCAACCGGAAAC 120
DB	69 TCAGGTCACCAACCAGCTATTGAACTTAATGTGAAAACATGATGCCAACCGGAAAC 128
OY	121 CCTATCCCCGACAGCCCACTGTGTCCCACTGTCTACAGAAGTGATCCGCTCAGTAC 180
DB	129 CCTATCCCCGACAGCCCACTGTGTCCCACTGTCTACAGAAGTGATCCGCTCAGTAC 188
OY	181 TACCCGTCCTCCGTCGCCAAGTAGCGCCCGAGAGGCTCTGACGCAGAGCTTCCAACCCGTC 240
DB	189 TACCCGTCCTCCGTCGCCAAGTAGCGCCCGAGAGGCTCTGACGCAGAGCTTCCAACCCGTC 248
OY	241 GTCTGACGACGAGCCCAATCCCATCC 267
DB	249 GTCTGACGACGAGCCCAATCCCATCC 275

XX AC AAD13168;
 XX XX 16-OCT-2001 (first entry)
 DT XX
 DE Human serine protease DNA.
 XX
 KM Human; transmembrane serine protease; membrane-type serine protease;
 KM MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;
 KM lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;
 KM serine protease; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 57..1535
 FT /tag= a
 FT /product= "Human serine protease"
 XX
 PN W0200157194-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 02-FEB-2001; 2001MO-US03471.
 XX
 PR 03-FEB-2000; 2000US-0179982.
 PR 18-FEB-2000; 2000US-0183542.
 PR 22-JUN-2000; 2000US-0213124.
 PR 26-JUL-2000; 2000US-0220970.
 PR 08-SEP-2000; 2000US-0657986.
 PR 22-SEP-2000; 2000US-0234840.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Madison EL, Ong EO, Yeh J;
 XX
 DR WPI; 2001-488877/53.
 DR P-PSDB; AAE06943.
 XX
 PT Novel single chain polypeptide comprising protease domain of type-II
 PT membrane-type serine protease or its catalytically active portion
 PT useful for treating and preventing cancer and tumor
 XX
 PS Disclosure; Page 250-252; 256pp; English.
 XX
 CC The invention relates to transmembrane serine proteases and their
 CC corresponding nucleotides and the protease domain of a type-II
 CC membrane-type serine protease (MTSP). MTSP is useful for identifying
 CC compounds that modulate or inhibit its proteolytic activity and for
 CC formulating a medicament for treating neoplastic disease. MTSP and
 CC its corresponding nucleotides are useful in preventing or treating
 CC tumours or cancers such as lung carcinoma, colon adenocarcinoma and
 CC ovarian carcinoma, in diagnostics and in hybridisation assays. MTSP
 CC is useful as a diagnostic marker for tumour development, growth and/or
 CC progression and as immunogens to generate antibodies that specifically
 CC bind to it. MTSP DNA is useful in a yeast two-hybrid system and in
 CC gene therapy. The present sequence is a DNA encoding human serine
 CC protease protein.
 CC
 SO Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
 XX
 QY Query Match 100.0%; Score 267; DB 22; Length 2479;
 DB Best Local Similarity 100.0%; Pred. No. 5,2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGATACCTATCTACTGATGCTGTGATAAGCAGATGGCTTGAAC 60
 DB 9 GAACATTCGATACCTATCTACTGATGCTGTGATAAGCAGATGGCTTGAAC 68
 QY 61 TCAGGCTCACACACAGCTATGACCTTACTATGAAAACATGATCAACCCGGAAC 120
 DB 69 TCAGGCTCACACACAGCTATGACCTTACTATGAAAACATGATCAACCCGGAAC 128

QY 121 CCCTATCCCGCAGACGCCACATGATGCTCCCACTGTCTACAGAGGATCCGGCTCACTAC 180
 DB 129 CCCTATCCCGCAGACGCCACATGATGCTCCCACTGTCTACAGAGGATCCGGCTCACTAC 188
 QY 181 TACCCGTCCTCCCGTCCCGCAGTACGCCCGCAGAGGTCTCTGACGACGCTTCCAAACCCGTC 240
 DB 189 TACCCGTCCTCCCGTCCCGCAGTACGCCCGCAGAGGTCTCTGACGACGCTTCCAAACCCGTC 248
 QY 241 GTCTGACAGCAGCCCAATCCCATCC 267
 DB 249 GTCTGACAGCAGCCCAATCCCATCC 275
 XX
 RESULT 11
 AAH93928
 ID AAH93928 standard; cDNA; 2479 BP.
 XX
 AC AAH93928;
 XX
 DT 04-OCT-2001 (first entry)
 XX
 DE Human transmembrane protease serine 2 full length cDNA sequence.
 XX
 KM Human; prostate cancer; prostate-specific; diagnosis; vaccine;
 KM cytostatic; gene therapy; metastasis; ss.
 OS Homo sapiens.
 XX
 PN W0200151633-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001MO-US01574.
 XX
 PR 14-JAN-2000; 2000US-0483672.
 XX
 PA (CORV-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kaloje MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skelky YAW,
 PI Wang A, Meagher MJ;
 XX
 DR WPI; 2001-425873/45.
 XX
 PT New polynucleotide encoding a prostate-specific protein, for
 PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines
 XX
 PS Claim 1; Page 518-519; 543pp; English.
 XX
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II), (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93157 to AAH93944 and AAH01115 to
 CC AAH01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.
 CC
 SO Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
 XX
 QY Query Match 100.0%; Score 267; DB 22; Length 2479;
 DB Best Local Similarity 100.0%; Pred. No. 5,2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGATACCTATCTACTGATGCTGTGATAAGCAGATGGCTTGAAC 60

Db 9 GAACATTCCAGATACCTATCTATCTGATGATACGCAAGATGCTTTGAAC 68
 Qy 61 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
 Db 69 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
 Qy 121 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 180
 Db 129 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 188
 Qy 181 TACCCGTCGCCCGTGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCCGTCGCCCGTGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 248
 Qy 241 GTCTGCACGACGCCCAATCCCATCC 267
 Db 249 GTCTGCACGACGCCCAATCCCATCC 275

RESULT 12

ABR92201
 ID ABR92201 standard; DNA; 2479 BP.

AC ABR92201;

DT 15-AUG-2002 (first entry)

DE Prostate cancer-associated DNA sequence #87.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
 XX gene therapy; gene; ds.

OS Mammalia.

PN W0200230268-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US32045.

XX 13-OCT-2000; 2000US-0687576.

XX 08-DEC-2000; 2000US-0733288.

XX 08-DEC-2000; 2000US-0733288.

XX 24-JAN-2001; 2001US-263957P.

XX 16-MAR-2001; 2001US-276791P.

XX 16-MAR-2001; 2001US-276888P.

XX 06-APR-2001; 2001US-281922P.

XX 24-APR-2001; 2001US-286214P.

XX 30-APR-2001; 2001US-0847046.

XX 04-MAY-2001; 2001US-288589P.

XX (BOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezl P,

XX WPI; 2002-471335/50.

XX P-PSDB; ABG61885.

XX Detecting a prostate cancer-associated transcript in a cell in a
 XX patient, useful for diagnosing prostate cancer (PC) or screening
 XX modulators of PC, by determining if prostate cancer-associated genes
 XX are expressed in a prostate tissue -

XX Claim 22; Page 371-372; 436pp; English.

XX The present invention relates to methods of detecting a prostate
 XX cancer-associated transcript in a cell from a patient. The method
 XX comprises contacting a biological sample from the patient with
 XX prostate cancer-associated polynucleotides (designated PC genes) that
 XX selectively hybridize to a sequence that is at least 80% identical
 XX to them. The prostate cancer-associated polynucleotide sequences
 XX are differentially expressed in prostate tumour tissue or in
 XX prostate cancer and are derived from the tissues of various

CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
 CC The methods of the invention are useful for diagnosing and treating
 CC prostate cancer in mammals. The prostate cancer-associated genes are
 CC useful for diagnosing or treating prostate cancer, as well as for
 CC identifying modulators of prostate cancer or agents that inhibit
 CC prostate cancer. The nucleic acid sequences are particularly useful
 CC in gene therapy, as a vaccine or in antisense applications.
 CC ABR92115-ABR92263 represent prostate cancer-associated polynucleotide
 CC sequences.

XX Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other.

XX Query Match 100.0%; Score 267; DB 24; Length 2479;

XX Best Local Similarity 100.0%; Pred. No. 5,2e-73;

XX Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACATTCCAGATACCTATCTATCTGATGATACGCAAGATGCTTTGAAC 60
 Db 9 GAACATTCCAGATACCTATCTATCTGATGATACGCAAGATGCTTTGAAC 68
 Qy 61 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
 Db 69 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
 Qy 121 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 180
 Db 129 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 188
 Qy 181 TACCCGTCGCCCGTGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCCGTCGCCCGTGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 248
 Qy 241 GTCTGCACGACGCCCAATCCCATCC 267
 Db 249 GTCTGCACGACGCCCAATCCCATCC 275

RESULT 13

ABL95535
 ID ABL95535 standard; cDNA; 2479 BP.

AC ABL95535;

DT 19-JUL-2002 (first entry)

XX Human transmembrane protease serine 2 cDNA sequence SEQ ID NO 894.

XX Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
 XX gene therapy; gene; ss.

XX Homo sapiens.

XX US2002022248-A1.

XX 21-FEB-2002.

XX 12-JAN-2001; 2001US-0759143.

XX 25-FEB-1997; 97US-0806099.

XX 01-AUG-1997; 97US-0904804.

XX 09-FEB-1998; 98US-0020956.

XX 25-FEB-1998; 98US-0030607.

XX 14-JUL-1998; 98US-0115453.

XX 23-SEP-1998; 98US-0159812.

XX 15-JAN-1999; 99US-0232149.

XX 09-APR-1999; 99US-0288946.

XX 13-JUL-1999; 99US-0352616.

XX 18-NOV-1999; 99US-0439313.

XX 14-JAN-2000; 2000US-0483672.

XX 27-MAR-2000; 2000US-0536857.

XX 09-MAY-2000; 2000US-0568100.

XX 12-MAY-2000; 2000US-0570737.

PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605793.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX

PA (XUJ/) XU J.
 PA (DILL/) DILLON D. C.
 PA (MITC/) MITCHAM J. L.
 PA (HARL/) HARLOCKER S. L.
 PA (JIAN/) JIANG Y.
 PA (KALO/) KALOS M. D.
 PA (PANG/) PANGER G. R.
 PA (RETT/) RETTER M. W.
 PA (STOL/) STOLK J. A.
 PA (DAYC/) DAY C. H.
 PA (VEDV/) VEDVICK T. S.
 PA (CART/) CARTER D.
 PA (LISX/) LI S. X.
 PA (WANG/) WANG A.
 PA (SKEL/) SKEIKY Y. A. W.
 PA (HEPL/) HEPLER W. T.
 PA (HEND/) HENDERSON R. A.

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD,
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 DR WPI; 2002-255649/30.
 XX

PT New prostate-specific polynucleotides for diagnosing and treating
 PT diseases, in particular prostate cancer, and as markers for the
 PT progression of cancer
 XX

XX Claim 1; SEQ ID NO 894; 87pp; English.

CC The present invention provides prostate-specific coding sequences and
 CC their encoded proteins. These can be used in the diagnosis and treatment
 CC of cancer, particularly prostate cancer. The present sequence is a CDNA
 CC described in the invention.
 XX

SO Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;

Query Match 100.0%; Score 267; DB 24; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 5.2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATAACAGCAGATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTATCATCTGATGCTGTGATAACAGCAGATGGCTTTGAAC 68
 QY 61 TAGAGGTACACCACTTGTGACTTACTAATAAACAATGATTCACACCGAAAC 120
 DB 69 TAGAGGTACACCACTTGTGACTTACTAATAAACAATGATTCACACCGAAAC 128
 QY 121 CCTATTCCTGACAGCCCACTGTGTCTCCCACTGTCTACAGAGGATCGGCTCAGTAC 180
 DB 129 CCTATTCCTGACAGCCCACTGTGTCTCCCACTGTCTACAGAGGATCGGCTCAGTAC 188
 QY 181 TACCGTCCCGGTGCTGACAGTACGCGCCGAGGGCTCTGACGAGGCTTCAACCCCGTC 240
 DB 189 TACCGTCCCGGTGCTGACAGTACGCGCCGAGGGCTCTGACGAGGCTTCAACCCCGTC 248
 QY 241 GTCGTGACGACGACCAAAATCCCATCC 267
 DB 249 GTCGTGACGACGACCAAAATCCCATCC 275

RESULT 14
 AAD28779
 ID AAD28779 standard; DNA; 2479 BP.

XX AAD28779;
 AC
 XX
 DT 07-MAY-2002 (first entry)
 XX
 DB Human TMPRSS2 gene.
 XX

KW Serine protease; 20P1F12/TMPRSS2; 20P1F12-GTC1; cell growth; neoplasia;
 KW cancer; vaccine; human; TMPRSS2 gene; de.
 XX

OS Homo sapiens.

FX Key Location/Qualifiers
 FH 57..1535
 FT /*tag= a
 FT /product= "Human TMPRSS2 protein"
 FT

PN W0200204953-A2.

PD 17-JAN-2002.

PF 12-JUL-2001; 2001MO-US22168.

PR 12-JUL-2000; 2000US-0615285.

PA (AGEN-) AGENSYS INC.

PI Saferan D, Rattano AB, Hubert RS, Jakobovits A, Paris M,
 PI Challita-Eld PM;
 PI

DR WPI; 2002-154967/20.
 DR P-SDB; AAE18097.

PT Examining a biological sample for evidence of dysregulated cellular
 PT growth, comprises comparing the status of prostate-specific,
 PT androgen-regulated, secreted serine protease, 20P1F12/TMPRSS2, in a
 PT corresponding normal sample
 XX

PS Example 3; Fig 2; 161pp; English.

CC The present invention relates to methods and compositions for the
 CC diagnosis and therapy of prostate, colon, bladder, lung, ovarian and
 CC kidney cancer derived from or based on a normally prostate-specific,
 CC androgen regulated, cell membrane associated secreted serine protease
 CC termed 20P1F12/TMPRSS2. The invention further relates to a method of
 CC examining a biological sample for evidence of dysregulated cellular
 CC growth comprising comparing the status of 20P1F12/TMPRSS2 gene (also
 CC designated 20P1F12-GTC1) in the sample to the status of 20P1F12/TMPRSS2
 CC in a corresponding normal sample. The invention also relates to 20P1F12/
 CC TMPRSS2 polynucleotides and their corresponding proteins. Methods of the
 CC invention are used for examining a sample such as blood, serum, stool,
 CC urine, semen, or biopsy tissue for evidence of dysregulated cell growth.
 CC The dysregulated cell growth is indicative of bladder cancer, lung
 CC cancer, kidney cancer or ovarian cancer. It is useful for identifying
 CC evidence of a neoplasm in a sample. Vaccines comprising an immunogenic
 CC portion of 20P1F12/TMPRSS2 are useful for inhibiting growth of a cell
 CC expressing 20P1F12/TMPRSS2 in a patient suffering from bladder cancer,
 CC lung cancer, ovarian cancer or metastatic cancer. The present sequence
 CC is human TMPRSS2 gene.
 XX

SO Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;

Query Match 100.0%; Score 267; DB 24; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 5.2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATAACAGCAGATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTATCATCTGATGCTGTGATAACAGCAGATGGCTTTGAAC 68
 QY 61 TAGAGGTACACCACTTGTGACTTACTAATAAACAATGATTCACACCGAAAC 120
 DB 69 TAGAGGTACACCACTTGTGACTTACTAATAAACAATGATTCACACCGAAAC 128

QY 121 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGTGCATCCGGCTCAGTAC 180
 DB 129 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGTGCATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGCTTCCAAACCCCGTC 240
 DB 189 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGCTTCCAAACCCCGTC 248
 QY 241 GTCTGCACGACGCCCAATATCCCATCC 267
 DB 249 GTCTGCACGACGCCCAATATCCCATCC 275

RESULT 15

AAA08803
 ID AAA08803 standard; cDNA; 3245 BP.

AC AAA08803;

DT 01-AUG-2000 (first entry)

DE Androgen-inducible gene clone HrpCa6/7 found in LNCaP cells.

KW Androgen inducible; testosterone; prostate cancer; cytostatic;

KM TMPRSS2; diagnosis; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 145..1623
 /tag= a
 /product= HrpCa6-7

PN WO200018961-A2.

PD 06-APR-2000.

PF 30-SEP-1999; 99WO-US22535.

PR 30-SEP-1998; 98US-0163759.

PR 30-SEP-1998; 98US-0164159.

PA (MILL-) MILLENNIUM PHARM INC.

PI Macbeth KJ, Shyjan AW,

DR MPI; 2000-293182/25.

DR P-PSDB; AAY92050.

PT Novel methods for identifying compounds for treating prostate cancer
 PT comprising measuring the level of expression or activity of 1 or more
 PT of 11 genes or their products

PS Claim 1; Fig 3; 108pp; English.

XX AAA08801-09 are genes which are androgen (e.g. testosterone) inducible
 CC in androgen-dependent prostate cancer cells (e.g. LNCaP cells) and
 CC constitutively expressed in androgen-independent prostate cancer cells
 CC (e.g. LNCaP cells). Agents which decrease the expression or
 CC activity of these clones may slow or arrest the growth of prostate cancer
 CC cells or may kill them. HrpCa6/7 can be obtained from the sequence of
 CC the known gene for TMPRSS2. A compound useful for treating prostate
 CC cancer can be identified in a novel method comprising measuring the
 CC expression level, or activity, of HrpCa2, 3, 6/7, 8, 9, 10, 13, 14, 15,
 CC 19, or peripheral-type benzodiazepine receptor (PBR) in a cell, in the
 CC presence and absence of a test compound. The sequences may also be used
 CC in diagnosis of prostate cancer and to determine efficacy of treatment
 CC for prostate cancer.

XX Sequence 3245 BP; 774 A; 818 C; 870 G; 783 T; 0 other;

XX Query Match

100.0%; Score 267; DB 21; Length 3245;

Best Local Similarity 100.0%; Pred. No. 5.8e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGAGTACTTATCTATCTGATGCTGTTGATTAACGAGAATGCTTTGAAAC 60
 DB 97 GAACATTCGAGTACTTATCTATCTGATGCTGTTGATTAACGAGAATGCTTTGAAAC 156
 QY 61 TCAGGGTACCAACCAAGTATTGACCTTACTATGAAAACCATGATATCCACCGGAAAAC 120
 DB 157 TCAGGGTACCAACCAAGTATTGACCTTACTATGAAAACCATGATATCCACCGGAAAAC 216
 QY 121 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGTGCATCCGGCTCAGTAC 180
 DB 217 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGTGCATCCGGCTCAGTAC 276
 QY 181 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGCTTCCAAACCCCGTC 240
 DB 277 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGCTTCCAAACCCCGTC 336
 QY 241 GTCTGCACGACGCCCAATATCCCATCC 267
 DB 337 GTCTGCACGACGCCCAATATCCCATCC 363

Search completed: January 31, 2003, 06:53:12
 Job time : 312 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 31, 2003, 06:47:46 ; Search time 2245 Seconds

(Without alignments)
1926.145 Million cell updates/sec

Title: US-09-807-201-7

Sequence: 1 gaacattccagataccatc.....cgccagcccaatcccatcc 267

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estro.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_hiv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_oth.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	267	100.0	866 13	BI761763 603046751
2	209.8	78.6	300 9	AU099721 AU099721
3	127.6	47.8	514 10	AW044734 AW044734
4	127.2	47.6	345 9	AI613666 AI613666
5	127.2	47.6	398 9	AA451050 AA451050
6	127.2	47.6	456 9	AA832852 AA832852

7	127.2	47.6	489 10	AW912791
8	127.2	47.6	555 10	AW211403
9	127.2	47.6	583 10	BQ108875
10	127.2	47.6	723 13	BI850775
11	127.2	47.6	783 12	BG176274
12	127.2	47.6	838 13	BI144021
13	127.2	47.6	915 13	BI143513
14	127.2	47.6	944 14	BQ228388
15	127.2	47.6	1019 13	BE916909
16	127.2	47.6	1137 13	BI687501
17	125.6	47.0	756 13	BI556978
18	125.6	47.0	796 13	BI556978
19	125.6	47.0	921 13	BI144727
20	118.8	44.5	741 13	BI683800
21	117.8	44.1	846 12	BG870706
22	116.8	43.7	486 10	AW909202
23	112	41.9	790 12	BF144549
24	110.2	41.3	338 12	BF001767
25	105.4	39.5	497 12	BR857022
26	92.2	34.5	633 12	BR916896
27	86.8	32.5	934 12	BF788678
28	82.6	30.9	177 12	BE696955
29	79.4	29.7	523 12	BF190896
30	77.2	28.9	599 13	BI105806
31	69	25.8	782 12	BR675039
32	69	25.8	1734 11	BC015819
33	65.6	24.6	795 13	BG962521
34	64	24.0	321 12	BF056592
35	64	24.0	618 10	AW604314
36	64	24.0	676 10	AW376911
37	56	21.0	221 10	BB870649
38	51.8	19.4	508 12	BE681173
39	45.6	17.1	703 10	BE376601
40	39.2	14.7	154 10	BB602684
41	39.2	14.7	909 9	AF122350
42	38.6	14.5	843 10	BE620434
43	38	14.2	848 17	AG600231
44	37.8	14.2	962 14	BQ715655
45	37.6	14.1	400 17	CNS03Y9R

ALIGNMENTS

RESULT 1
BI761763
LOCUS 866 bp mRNA linear EST 25-SEP-2001
DEFINITION 603046751F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5187060 5',
BI761763
VERSION BI761763.1 GI:15753341
ACCESSION
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 866)
NIH-MGC http://mgs.cbl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium information can be
http://image.llnl.gov
Plate: L14M11467 row: e column: 13
High quality sequence stop: 781.
Location/Qualifiers
1. 866

FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5187060"
/clone_1ib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pcmv-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dt primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

BASE COUNT      204 a      251 c      236 g      175 t
ORIGIN

Query Match      100.0%; Score 267; DB 13; Length 866;
Best Local Similarity 100.0%; Pred. No. 2.1e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 60
    |||
DB 70 GAACATTCCAGATACCTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 129
    |||

QY 61 TCAGGGTCCACACAGTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 120
    |||
DB 130 TCAGGGTCCACACAGTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 189
    |||

QY 121 CCTATCCGCAAGCCCACTGTGTGCTTCAAGAGGTGATCCGGCTCAGTAC 180
    |||
DB 190 CCTATCCGCAAGCCCACTGTGTGCTTCAAGAGGTGATCCGGCTCAGTAC 249
    |||

QY 181 TACCCGTCCTCCGTCGTCAGTACGCCCGAGGGTCTCTGACGAGGCTTCAACCCCGTC 240
    |||
DB 250 TACCCGTCCTCCGTCGTCAGTACGCCCGAGGGTCTCTGACGAGGCTTCAACCCCGTC 309
    |||

QY 241 GTCTGCACGACGCCCAATCCCATCC 267
    |||
DB 310 GTCTGCACGACGCCCAATCCCATCC 336
    |||

RESULT 2
A0099721      300 bp      mRNA      linear      EST 05-APR-2001
LOCUS
DEFINITION
A0099721 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
KAT01641 similar to Human serine protease mRNA, mRNA sequence.
ACCESSION
A0099721
VERSION
A0099721.1 GI:13550850
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 300)
Suzuki,Y., Tsunoda,T., Taiba,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Oca,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.
In silico mapping of the 5'-ends of human mRNAs using full-length
enriched and 5'-end enriched cDNA libraries constructed by
Oligo-capping method
Unpublished (2001)
CONTACT: Yutaka Suzuki
Department of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
Email: ysuzuki@ms.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
Location/Qualifiers

```

```

source      1..300
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT01641"
/clone_1ib="Sugano Homo sapiens cDNA library"
/notes="Differential display comparison of untreated and
dimethylfumarate treated U937 cells"

BASE COUNT      67 a      89 c      88 g      56 t
ORIGIN

Query Match      78.6%; Score 209.8; DB 9; Length 300;
Best Local Similarity 99.1%; Pred. No. 8.8e-48;
Matches 211; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 60
    |||
DB 88 GAACATTCCAGATACCTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 147
    |||

QY 61 TCAGGGTCCACACAGTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 120
    |||
DB 148 TCAGGGTCCACACAGTATGATGCTGTGTGTAAGACAGATGGCTTGAAC 207
    |||

QY 121 CCTATCCGCAAGCCCACTGTGTGCTTCAAGAGGTGATCCGGCTCAGTAC 180
    |||
DB 208 CCTATCCGCAAGCCCACTGTGTGCTTCAAGAGGTGATCCGGCTCAGTAC 267
    |||

QY 181 TACCCGTCCTCCGTCGTCAGTACGCCCGAGG 213
    |||
DB 268 TACCCGTCCTCCGTCGTCAGTACGCCCGAGG 300
    |||

RESULT 3
A0044734      514 bp      mRNA      linear      EST 18-SEP-1999
LOCUS
DEFINITION
A0044734 um13h07.y1 Sugano mouse kidney mRna Mus musculus cDNA clone
IMAGE:2192221 5' similar to SW:TM52_HUMAN O15393 TRANSMEMBRANE
PROTEASE, SERINE 2 ;, mRNA sequence.
ACCESSION
A0044734
VERSION
A0044734.1 GI:5905331
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 514)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Marras,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y.,
B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurr,R., Ritter
,E., Kohn,S., Shin,T., Jackson,I., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
CONTACT: Marra M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1004673
Seq primer: custom primer used
High quality sequence stop: 482.
Location/Qualifiers
1..514
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2192221"
/clone_1ib="Sugano mouse kidney mRna"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
FEATURES
source

```

```

/tissue type="mammary gland"
/dev stage="4 weeks"
/lab host="DHI08"
/notes="Oxran; mammary gland; Vector: pT73D-Pac (Pharmacia
) with a modified polylinker; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TGTATCCACATCTGAAGTGGAGCGCGCGCAATGGTGTGTGTGTGTGTGTGT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
cDNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Feltina

```

a	109 c	84 g	60 c
---	-------	------	------

Query Match	47.6%;	Score 127.2;	DB 9;	Length 345;
Best Local Similarity	71.2%;	Pred. No. 7.1e-25;		
Matches 168; Conservative	0;	Mismatches 68;	Indels 0;	Gaps 0;

07 ATGCTGTGATTAACAGCAAGATGGCTTGAATCTCAGGGTTCACACCAAGCTATTTGACCTT 88
 29 ATGCTGTGATTAACAGCAAGATGGCTTGAATCTCAGGGTTCACACCAAGCTATTTGACCTT 88
 Db 82 ATGCTGTGATTAACAGCAAGATGGCTTGAATCTCAGGGTTCACCTCCAGGATTCGACCTT 141

89 ACTATGAAACCATGATACCAACCGGAAACCCCTATCCCGCAGAGCCCATGTGTCTCC 148

149	CCACTGTCTACGAGGTGCATCCGGCTCAGTACTACCCGATCCCGGCCCCAGTACGCC	208
Dy		
Db	142 GCTATGAGACCAACGGGATCAGTCTGAGGACATCTCTCTCCGAGACCAACAGTGGCTC	201

Db 202 CCATGGCTACAACTGTATTCAGCCAGTACTACCATTCCTCCAGTGGCTCAGTATGCTC 261

Db
262 CGAGGATTACAACGCAAGCCCTCAACATCTGTATCATCCACACACATCCCAAGTCTCA 317

[illegible]

DEFINITION	
V66G09:R1	solar_mammary_gland_NDMMG
IMAGE:	R850720_5', mRNA sequence.
AA451050	
AA451050	CT-31564720
REPERSON	
AA451050	1
MUS	MUSCULUS
CUDA	CUDA
CLONE	CLONE

	ASIMUS	Est.	house mouse.	
SOURCE			Mus musculus	
ORGANISM				

REFERENCE
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 398)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Thelings, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterson, R.

COMMENT
Contact: Maira M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine

Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LINT ; contact the

This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:502872

```
source 1. .398
        /organism="Mus musculus"
```

/db_xref="taxon:10090"
/clone="IMAGE:850720"
/clone_1ib="Soares_mammary_gland_NBMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5']
TGTACCAATCTGAAGTGGAGCGCGCGGAATGTTTTTTTTTTTTTTTT
T 3'; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

BASE COUNT 102 a 125 c 96 g 75 t

ORIGIN

Query Match 47.6%; Score 127.2; DB 9; Length 398;
Best Local Similarity 71.2%; Pred. No. 7.6e-25;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACGAGAGATGGCTTTGAATCAGGGTCACACGCTATTGACCTT 88
DB 81 ATGCTGTGACACGAGAGATGGCTTTGAATCAGGGTCACACGCTATTGACCTT 140

QY 89 ACTATGAAAACCATGATATCAACCGAATCCCTATCCCGACAGCCCACTGTGTC 148
DB 141 GCTATGAAACACGAGTATCAGTGTGACACATCTGTCTCCGACACACAGTGGCTC 200

QY 149 CCACTGTCTACGAGTGCATCCGCTCACTACCTCCGCTCCCGCCAGTACGCC 208
DB 201 CCAATGGCTCAACTGTATTCAGCCCACTACCTCACTCACTCACTCACTCACTC 260

QY 209 CGAGGGTCTGACGAGGCTTCCACCCGCTGTGACGACGACCAATCCCA 264
DB 261 CGAGGATTAACGAGGCTTCAACATCTGTATCCACACATCCCAAGTCTCA 316

RESULT 6
AA832852 456 bp mRNA linear EST 23-FEB-1998
LOCUS ub56g01.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:1381968 5', mRNA sequence.
ACCESSION AA832852
VERSION AA832852.1 GI:2906580
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (base 1 to 456)
REFERENCE
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le M., Martin, J., Morris, M.,
Schellberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lemon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:904436

Seq primer: -28m3 rev2 ET from Amerham
High quality sequence stop: 452.
Location/Qualifiers
1. 456
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1381968"
/clone_1ib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pRT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pRT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 99 a 138 c 114 g 105 t

ORIGIN

Query Match 47.6%; Score 127.2; DB 9; Length 456;
Best Local Similarity 71.2%; Pred. No. 8.1e-25;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACGAGAGATGGCTTTGAATCAGGGTCACACGCTATTGACCTT 88
DB 27 ATGCTGTGACACGAGAGATGGCTTTGAATCAGGGTCACACGCTATTGACCTT 86

QY 89 ACTATGAAAACCATGATATCAACCGAATCCCTATCCCGACAGCCCACTGTGTC 148
DB 87 GCTATGAAACACGAGTATCAGTGTGACACATCTGTCTCCGACACACAGTGGCTC 146

QY 149 CCACTGTCTACGAGTGCATCCGCTCACTACCTCCGCTCCCGCCAGTACGCC 208
DB 147 CCAATGGCTCAACTGTATTCAGCCCACTACCTCACTCACTCACTCACTCACTC 206

QY 209 CGAGGGTCTGACGAGGCTTCCACCCGCTGTGACGACGACCAATCCCA 264
DB 207 CGAGGATTAACGAGGCTTCAACATCTGTATCCACACATCCCAAGTCTCA 262

RESULT 7
AW912791 489 bp mRNA linear EST 25-MAY-2000
LOCUS uf45b02.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:1514283 5' similar to SW:TMS2_HUMAN O15393 TRANSMEMBRANE
PROTEASE, SERINE 2 /, mRNA sequence.
ACCESSION AW912791
VERSION AW912791.1 GI:8078415
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (base 1 to 489)
REFERENCE
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: uf45b02.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:941135
Seq primer: -40RP from G1bco
High quality sequence stop: 443.
Location/Qualifiers
1. 489
/organism="Mus musculus"
/db_xref="taxon:10090"

FEATURES
source

/clone="IMAGE:1514283"
 /clone_id="Soares mammary_gland_NMLMG"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="MDH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pT733 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Patricia Bonaldo."

BASE COUNT 117 a 151 c 116 g 103 t 2 others
 ORIGIN

Query Match 47.6%; Score 127.2; DB 10; Length 489;
 Best Local Similarity 71.2%; Pred. No. 8.4e-25;
 Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

29 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCAACCAAGTATGGACCTT 88
 Db ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCAACCAAGTATGGACCTT 137
 78 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCAACCAAGTATGGACCTT 137
 89 ACTATGAAAACCATGATATCAACCGGAAAACCCCTATCCGCAAGCCCACTGTGCTC 148
 Db GCTATGAAACCAAGGATATCATGTGAGCAATCTGTCTCCAGAGACACCAAGTGGCTC 197
 138 GCTATGAAACCAAGGATATCATGTGAGCAATCTGTCTCCAGAGACACCAAGTGGCTC 197
 149 CCACTGTCTAAGAGTGCATCCGCTCAGTACTACCCGCTCCCGTCCCAAGTACGCCC 208
 Db CCAATGGCTCAACTTGTATCAAGCCCAAGTACTACCCGCTCCCGTCCCAAGTACGCCC 257
 198 CCAATGGCTCAACTTGTATCAAGCCCAAGTACTACCCGCTCCCGTCCCAAGTACGCCC 257
 209 CGAGGGTCTGACGCAAGCTTCAACCCCGCTGTGTGACGAGCCCAAAATCCCA 264
 Db 258 CGAGGATTAACAGCAAGCTTCAATCTGTATTCACACATCCCAAGTCTCA 313

RESULT 8
 LOCUS AW211403 555 bp mRNA linear EST 03-DEC-1999
 DEFINITION ubo0e12.y1 NCI CGAP Mam3 Mus musculus cDNA clone IMAGE:2648878 5'
 similar to SW:TW52_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;
 mRNA sequence.
 ACCESSION AW211403
 VERSION AW211403.1 GI:6517358
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 555)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Lohar Hemmighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdtp/image/image.html

FEATURES
 source
 MGI:1029330
 Seg primer: -40RP from Gibco
 High quality sequence stop: 427.
 Location/Qualifiers
 1..555
 /organism="Mus musculus"
 /strain="129, C57BL/6J, FVB/N"
 /db_xref="taxon:10090"

/clone="IMAGE:2648878"
 /clone_id="NCI CGAP Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="MDH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lohar Hemmighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 123 a 165 c 142 g 125 t
 ORIGIN

Query Match 47.6%; Score 127.2; DB 10; Length 555;
 Best Local Similarity 71.2%; Pred. No. 8.9e-25;
 Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

29 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCAACCAAGTATGGACCTT 88
 Db ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCAACCAAGTATGGACCTT 118
 59 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCAACCAAGTATGGACCTT 118
 89 ACTATGAAAACCATGATATCAACCGGAAAACCCCTATCCGCAAGCCCACTGTGCTC 148
 Db GCTATGAAACCAAGGATATCATGTGAGCAATCTGTCTCCAGAGACACCAAGTGGCTC 178
 119 GCTATGAAACCAAGGATATCATGTGAGCAATCTGTCTCCAGAGACACCAAGTGGCTC 178
 149 CCACTGTCTAAGAGTGCATCCGCTCAGTACTACCCGCTCCCGTCCCAAGTACGCCC 208
 Db CCAATGGCTCAACTTGTATCAAGCCCAAGTACTACCCGCTCCCGTCCCAAGTACGCCC 238
 179 CCAATGGCTCAACTTGTATCAAGCCCAAGTACTACCCGCTCCCGTCCCAAGTACGCCC 238
 209 CGAGGGTCTGACGCAAGCTTCAACCCCGCTGTGTGACGAGCCCAAAATCCCA 264
 Db 239 CGAGGATTAACAGCAAGCTTCAATCTGTATTCACACATCCCAAGTCTCA 294

RESULT 9
 LOCUS B0108875 583 bp mRNA linear EST 16-APR-2002
 DEFINITION imageqc.8.2001/bmp14bdr81.y1 NCI CGAP Mam6 Mus musculus cDNA
 clone IMAGE:5342443 5', mRNA sequence.
 ACCESSION B0108875
 VERSION B0108875.1 GI:20158529
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 583)
 Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
 Prange,C.K.
 The I.M.A.G.E. Consortium quality control effort: clone
 resequencing for verification
 JOURNAL Unpublished (2001)
 COMMENT Other ESTs: B1683800
 Contact: Prange CK
 The I.M.A.G.E. Consortium
 Lawrence Livermore National Laboratory
 Livermore, CA, USA
 Email: help@image.llnl.gov
 This read has been verified (found to hit its original self in the
 correct orientation), as part of the I.M.A.G.E. Consortium quality
 control effort. High quality sequence is defined as having 100 or
 more base pairs with a phred quality value of 20 or greater, where
 or greater marks the beginning and end of the sequence. For
 information on obtaining this clone, please contact
 info@image.llnl.gov.
 Plate: LLAM11869 row: o column: 20
 Seg primer: m3rpl
 High quality sequence stop: 583.
 Location/Qualifiers
 1..583
 /organism="Mus musculus"

FEATURES
 source
 High quality sequence stop: 583.
 Location/Qualifiers
 1..583
 /organism="Mus musculus"


```

/strain="FVB/N"
/db xref="taxon:10090"
/clone="IMAGE:5342443"
/clone_lib="NCI CGAP Mam6"
/sex="Female, virgin"
/issue_type="Infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      124 a 175 c 160 g 123 t
ORIGIN

```

```

Query Match
Best Local Similarity 71.2%; Pred. No. 9,2e-25;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

```

QY 29 ATGCTGTTGATTAACAGAGATGGCTTTGAATCAAGGTCACACAGCTATTGACCTT 88
DB 213 ATGCTGTCACACAGGAGATGGATTTGAATCAAGGTCACCTCCAGAAATGGACCTT 272
QY 89 ACTATGAAAACCATGTGATACCAACCGAAACCCCTATCCCGCACAGCCCACTGTGCTC 148
DB 273 GCTATGAGAACACAGGGATACATCTGTCTCTCCGAGACACAGTGGCTC 332
QY 149 CCACTGTCTACAGAGTGATCCGAGTCACTACTACCCGTCGCCCGCCAGTACGGCC 208
DB 333 CCAATGGCTTCAACTGTATTCAGCCAGTACTACCAATTCAGTGCCTCAAGTATGCTC 392
QY 209 CGAGGGTCTCTGACGAGGCTTCCAAACCCGTCGTCTGACAGCAGCCCAATCCCA 264
DB 393 CGAGGATTAACAAGCAGCTCAACATCTGTCTATCCACACATCCCAAGTCTCA 448

```

```

RESULT 10
BI850775      723 bp      mRNA      linear      EST 10-OCT-2001
LOCUS         imagegc.1.2001/smf23bdr81.y1 NCI CGAP Mam1 Mus musculus cDNA
DEFINITION    clone IMAGE:4461079 5', mRNA sequence.
ACCESSION     BI850775
VERSION       BI850775.1 GI:16004262
KEYWORDS      EST.
SOURCE        house mouse.
ORGANISM      Mus musculus

```

```

REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
              Prange, C.K.

```

```

JOURNAL       The I.M.A.G.E. Consortium quality control effort: clone
COMMENT        Resequencing for verification.
              Unpublished (2001)
              Other ESTs: BG176274
              Contact: Prange CK
              The I.M.A.G.E. Consortium
              Lawrence Livermore National Laboratory
              Livermore, CA, USA

```

```

Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the
correct orientation), as part of the I.M.A.G.E. Consortium quality
control effort. High quality sequence is defined as having 100 or
more base pairs with a phred quality value of 20 or greater, where
a sliding window of 4 base pairs marks the beginning and end of the
sequence. For information on obtaining this clone, please contact
info@image.llnl.gov.
Plate: LHAM10263 row: 1 column: 8
Seq primer: m33p1
High quality sequence stop: 723.
Location/Qualifiers
1..723
/organism="Mus musculus"

```

```

/strain="FVB/N"
/db xref="taxon:10090"
/clone="IMAGE:4461079"
/clone_lib="NCI CGAP Mam1"
/sex="Female, virgin"
/issue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      177 a 197 c 190 g 155 t
ORIGIN

```

```

Query Match
Best Local Similarity 71.2%; Pred. No. 1e-24;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

```

QY 29 ATGCTGTTGATTAACAGAGATGGCTTTGAATCAAGGTCACACAGCTATTGACCTT 88
DB 81 ATGCTGTCACACAGGAGATGGATTTGAATCAAGGTCACCTCCAGAAATGGACCTT 140
QY 89 ACTATGAAAACCATGTGATACCAACCGAAACCCCTATCCCGCACAGCCCACTGTGCTC 148
DB 141 GCTATGAGAACACAGGGATACATCTGTCTCTCCGAGACACAGTGGCTC 200
QY 149 CCACTGTCTACAGAGTGATCCGAGTCACTACTACCCGTCGCCCGCCAGTACGGCC 208
DB 201 CCAATGGCTTCAACTGTATTCAGCCAGTACTACCAATTCAGTGCCTCAAGTATGCTC 260
QY 209 CGAGGGTCTCTGACGAGGCTTCCAAACCCGTCGTCTGACAGCAGCCCAATCCCA 264
DB 261 CGAGGATTAACAAGCAGCTCAACATCTGTCTATCCACACATCCCAAGTCTCA 316

```

```

RESULT 11
BG176274      783 bp      mRNA      linear      EST 06-FEB-2001
LOCUS         602338030P1 NCI CGAP Mam1 Mus musculus cDNA clone IMAGE:4461079 5',
DEFINITION    mRNA sequence.
ACCESSION     BG176274
VERSION       BG176274.1 GI:12682977
KEYWORDS      EST.
SOURCE        house mouse.
ORGANISM      Mus musculus

```

```

REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         NIH-MGC http://mgc.nci.nih.gov/.

```

```

JOURNAL       National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT        Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-remail.nih.gov

```

```

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10263 row: 1 column: 08
High quality sequence stop: 665.
Location/Qualifiers
1..783
/organism="Mus musculus"
/strain="FVB/N"
/db xref="taxon:10090"
/clone="IMAGE:4461079"
/clone_lib="NCI CGAP Mam1"
/issue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI;

```


Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH

BASE COUNT

188 a 213 c 211 g 171 t

ORIGIN

Query Match 47.6%; Score 127.2; DB 12; Length 783;
Best Local Similarity 71.2%; Pred. No. 1.1e-24;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACGACAGATGGCTTTGAATCTGAGGCTCACACCAAGCTATTGACCTT 88
DB 81 ATGCTGTGACACAGGAGATGATGATTAATCTGAGGCTCACACCAAGCTATTGACCTT 140
QY 89 ACTATGAAAACCATGATATACCAACCGAAGAACCCCTATCCGCGACAGCCCACTGTGTCTC 148
DB 141 GCTATGAGAACCAACGGGATATCATGTCTGACACATCTGTCTCTCGAAGACCAACAGTGGCTC 200
QY 149 CCACTGTCTACAGAGGTGATCCGCTCAGTACTACCCGCTCCCGTCCCGCAGTACGCCC 208
DB 201 CCAATGCTTACCAACTTGTATCAGCCCACTACTACCAATCTCCAGTGGCTCAGTATGCTC 260
QY 209 CGAGGCTCTACAGAGGCTTCCACCCCGTGTCTGACAGGACCAATATCCCA 264
DB 261 CGAGGATTAACAACGACGCTCAACATCTGTATCCACACATCCCAAGTCTCA 316

RESULT 12

LOCUS B1144021 838 bp mRNA linear EST 05-JUL-2001
DEFINITION 60290812AF1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5065205
ACCESSION B1144021
VERSION B1144021.1 GI:14604022
KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 838)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LHAM1175 row: p column: 06
High quality sequence stop: 751.
Location/Qualifiers
1. 838

COMMENT

JOURNAL

TITLE

AUTHORS

REFERENCE

FEATURES

source

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5065205"
/lab_host="NCI CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library." |"
BASE COUNT 208 a 224 c 227 g 179 t
ORIGIN

Query Match 47.6%; Score 127.2; DB 13; Length 838;
Best Local Similarity 71.2%; Pred. No. 1.1e-24;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACGACAGATGGCTTTGAATCTGAGGCTCACACCAAGCTATTGACCTT 88
DB 87 ATGCTGTGACACAGGAGATGATGATTAATCTGAGGCTCACACCAAGCTATTGACCTT 146
QY 89 ACTATGAAAACCATGATATACCAACCGAAGAACCCCTATCCGCGACAGCCCACTGTGTCTC 148
DB 147 GCTATGAGAACCAACGGGATATCATGTCTGACACATCTGTCTCTCGAAGACCAACAGTGGCTC 206
QY 149 CCACTGTCTACAGAGGTGATCCGCTCAGTACTACCCGCTCCCGTCCCGCAGTACGCCC 208
DB 207 CCAATGCTTACCAACTTGTATCAGCCCACTACTACCAATCTCCAGTGGCTCAGTATGCTC 266
QY 209 CGAGGCTCTACAGAGGCTTCCACCCCGTGTCTGACAGGACCAATATCCCA 264
DB 267 CGAGGATTAACAACGACGCTCAACATCTGTATCCACACATCCCAAGTCTCA 322

RESULT 13

LOCUS B1143513 915 bp mRNA linear EST 05-JUL-2001
DEFINITION 602907696F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5064815
ACCESSION B1143513
VERSION B1143513
KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 915)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LHAM1174 row: o column: 24
High quality sequence start: 3
High quality sequence stop: 810.
Location/Qualifiers
1. 915

COMMENT

JOURNAL

TITLE

AUTHORS

REFERENCE

FEATURES

source

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5064815"
/lab_host="NCI CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library." |"
BASE COUNT 219 a 248 c 250 g 197 t
ORIGIN

Query Match 47.6%; Score 127.2; DB 13; Length 915;
Best Local Similarity 71.2%; Pred. No. 1.1e-24;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACGACAGATGGCTTTGAATCTGAGGCTCACACCAAGCTATTGACCTT 88
DB 162 ATGCTGTGACACAGGAGATGATGATTAATCTGAGGCTCACACCAAGCTATTGACCTT 221
QY 89 ACTATGAAAACCATGATATACCAACCGAAGAACCCCTATCCGCGACAGCCCACTGTGTCTC 148
DB 222 GCTATGAGAACCAACGGGATATCATGTCTGACACATCTGTCTCTCGAAGACCAACAGTGGCTC 281
QY 149 CCACTGTCTACAGAGGTGATCCGCTCAGTACTACCCGCTCCCGTCCCGCAGTACGCCC 208

Db 282 CCATGCTACACTTATTCACGCCAGTACTCCATCTCCAGTCCCTCAGTATCTC 341
 QY 209 CGAGGGTCTGACGACGAGCTTCCAAACCCGTCGTCTGACGACGACCCAAATCCCA 264
 Db 342 CGAGGATTACAGCAGCAGCTCAACATCTGTATTCACACACATCCCAAGTCTCA 397

RESULT 14
 BQ228388 944 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ228388
 DEFINITION AGENCOURT 7578382 NCI CGAP St1 Mus musculus cDNA clone
 IMAGE:6051691 5', mRNA sequence.
 ACCESSION BQ228388
 VERSION BQ228388.1 GI:20409788
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 944)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM1305 row: 0 column: 20
 High quality sequence stop: 707.

FEATURES
 source Location/Qualifiers
 1..944
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:6051691"
 /lab_host="NCI CGAP St1"
 /note="Organ: stomach; Vector: pCMV-Sport6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.77 kb. Library constructed by Life
 Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 231 a 243 c 263 g 207 t

Query Match 47.6%; Score 127.2; DB 14; Length 944;
 Best Local Similarity 71.2%; Pred. No. 1.2e-24;
 Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTATTAACAGCAAGATGGCTTGAATCAGGCTACCAACGATTTGACCTT 88
 Db 84 ATCTGTGTACACAGGAGATGGCATTTGAATCAAGGTGCTCCAGAAATGGACCTT 143
 QY 89 ACTATGAAACCATGATATCAACCGGAAACCCCTATCCCGACAGCCACTGTGTCC 148
 Db 144 GCTATGAGAACACGGGTATCAGTGTGAGCAATCTCTCTCGAGACCAACAGTGGCTC 203
 QY 149 CCACTGTCTACAGAGTGCATCCGGCTCAGTACTACCCGTCGCCCGGAGTACGCC 208
 Db 204 CCAATGGCTACAACTGTATTCAGCCCAAGTACCACTCCAGTCCCAATAGCTC 263
 QY 209 CGAGGGTCTGACGACGAGCTTCCAAACCCGTCGTCTGACGACGACCCAAATCCCA 264
 Db 264 CGAGGATTACAGCAGCAGCTCAACATCTGTATTCACACACATCCCAAGTCTCA 319

RESULT 15
 BE916909 1019 bp mRNA linear EST 29-SEP-2000
 LOCUS BE916909

DEFINITION 60166450P1 NCI CGAP Mam1 Mus musculus cDNA clone IMAGE:3966301 5',
 mRNA sequence.
 BE916909
 ACCESSION BE916909
 VERSION BE916909.1 GI:10418022
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1019)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM9139 row: d column: 14
 High quality sequence stop: 600.

FEATURES
 source Location/Qualifiers
 1..1019
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:3966301"
 /clone_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 281 a 318 c 261 g 159 t

Query Match 47.6%; Score 127.2; DB 12; Length 1019;
 Best Local Similarity 71.2%; Pred. No. 1.2e-24;
 Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTATTAACAGCAAGATGGCTTGAATCAGGCTACCAACGATTTGACCTT 88
 Db 59 ATGCTGTACACAGGAGATGGCATTTGAATCAAGGTGCTCCAGAAATGGACCTT 118
 QY 89 ACTATGAAACCATGATATCAACCGGAAACCCCTATCCCGACAGCCACTGTGTCC 148
 Db 119 GCTATGAGAACACGGGTATCAGTGTGAGCAATCTCTCTCGAGACCAACAGTGGCTC 178
 QY 149 CCACTGTCTACAGAGTGCATCCGGCTCAGTACTACCCGTCGCCCGGAGTACGCC 208
 Db 179 CCAATGGCTACAACTGTATTCAGCCCAAGTACCACTCCAGTCCCAATAGCTC 238
 QY 209 CGAGGGTCTGACGACGAGCTTCCAAACCCGTCGTCTGACGACGACCCAAATCCCA 264
 Db 239 CGAGGATTACAGCAGCAGCTCAACATCTGTATTCACACACATCCCAAGTCTCA 294

Search completed: January 31, 2003, 08:28:28
 Job time : 2251 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 31, 2003, 06:47:46 ; Search time 3216 Seconds

(without alignments)
2416.184 Million cell updates/sec

Title: US-09-807-201-7

Perfect score: 1
Sequence: 1 gacattccagataccatc.....cgccagcccaatcccatcc 267

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

```

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_vl:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match Length	DB ID	Description
1	267	100.0	683	AX201123
2	267	100.0	683	AX267922
3	267	100.0	1738	AX395316
4	267	100.0	1740	AF270487
5	267	100.0	2479	AX201121
6	267	100.0	2479	AX267965
7	267	100.0	2479	AX267920
8	267	100.0	2479	AX306771
9	267	100.0	2479	AX395318
10	267	100.0	2479	HSU75329
11	267	100.0	3226	AF329454
12	267	100.0	3245	AX201156
13	267	100.0	3245	AX267955
14	267	100.0	3966	AX041973
15	219	82.0	1476	AX267957
16	219	82.0	1476	AX201158
17	219	82.0	1479	AX201157
18	219	82.0	1479	AX267956
19	219	82.0	1479	AF133453
20	204	76.4	108927	AP001609
21	204	76.4	132110	AP001610
22	204	76.4	340000	HS21C085
23	178	66.7	196	AX106370
24	178	66.7	196	AX140651
25	178	66.7	196	AX200521
26	178	66.7	196	AX267177
27	127.2	47.6	1735	AF243500
28	127.2	47.6	1753	AF113596
29	127.2	47.6	1755	AF199362
30	118.4	44.3	1473	AB073550
31	103.4	38.7	220877	AC024957
32	64	24.0	2209	AK092554
33	48	18.0	111	AX155245
34	48	18.0	3660	HS2676404
35	38.8	14.5	158684	AC121032
36	38.6	14.5	10444	PV1422133
37	38.4	14.4	191325	AC125714
38	37.4	14.0	143798	AC012402
39	37.4	14.0	176802	AC078859
40	37.2	13.9	137390	AC113133
41	37.2	13.9	165725	AC024027
42	37.2	13.9	179676	AC015649
43	37	13.9	1641	RAB11A10
44	36.8	13.8	162336	AC021594
45	36.8	13.8	184510	AC079059

ALIGNMENTS

RESULT 1
AX201123
LOCUS AX201123
DEFINITION Sequence 753 from Patent WO0151633.
ACCESSION AX201123
VERSION AX201123.1 GI:15390897
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 683)
REFERENCE Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
Stolk,J.A., Shetty,Y.A., Wang,A. and Meagher,M.J.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0151633-A 753 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
SOURCE 1. 683
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 160 a 209 c 171 g 143 t
ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 683;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATAAGCAAGATGGCTTTGAAC 60
DB 9 GAACATTCAGATACCTATCTATCTGATGCTGTGATAAGCAAGATGGCTTTGAAC 68
QY 61 TCAGGGTCACCAACAGCTATTTGACCTTATGTAACCAAGTGAACCAACCGGAAAC 120
DB 69 TCAGGGTCACCAACAGCTATTTGACCTTATGTAACCAAGTGAACCAACCGGAAAC 128
QY 121 CCTATCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 180
DB 129 CCTATCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 188
QY 181 TACCGTCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 240
DB 189 TACCGTCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 248
QY 241 GTCTGACGAGCCCAATCCCATCC 267
DB 249 GTCTGACGAGCCCAATCCCATCC 275

RESULT 2
AX267922 683 bp DNA linear PAT 26-OCT-2001
LOCUS Sequence 896 from Patent WO0173032.
DEFINITION AX267922
ACCESSION AX267922.1 GI:16516527
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,
Kaijog, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,
Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Sketky, Y.A., Hepler, W.T.
and Henderson, R.A.
Compositions and methods for the therapy and diagnosis of prostate
cancer
Patent: WO 0173033-A 896 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
SOURCE 1. 683
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 160 a 209 c 171 g 143 t
ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 683;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATAAGCAAGATGGCTTTGAAC 60
DB 9 GAACATTCAGATACCTATCTATCTGATGCTGTGATAAGCAAGATGGCTTTGAAC 68
QY 61 TCAGGGTCACCAACAGCTATTTGACCTTATGTAACCAAGTGAACCAACCGGAAAC 120

DB 69 TCAGGGTCACCAACAGCTATTTGACCTTATGTAACCAAGTGAACCAACCGGAAAC 128
QY 121 CCTATCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 180
DB 129 CCTATCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 188
QY 181 TACCGTCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 240
DB 189 TACCGTCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 248
QY 241 GTCTGACGAGCCCAATCCCATCC 267
DB 249 GTCTGACGAGCCCAATCCCATCC 275

RESULT 3
AX395316 1738 bp DNA linear PAT 18-MAY-2002
LOCUS Sequence 1 from Patent WO0204953.
DEFINITION AX395316
ACCESSION AX395316.1 GI:21066313
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Safarian, D., Raitano, A.B., Hubert, R.S., Jakobovits, A., Faris, M. and
Challita-Eld, P.M.
Novel tumor antigen useful in diagnosis and therapy of bladder,
ovary, lung and kidney cancers
Patent: WO 0204953-A 1 17-JAN-2002;
Agensys, Inc. (US)
FEATURES Location/Qualifiers
SOURCE 1. 1738
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 433 a 459 c 473 g 373 t
ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 1738;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATAAGCAAGATGGCTTTGAAC 60
DB 64 GAACATTCAGATACCTATCTATCTGATGCTGTGATAAGCAAGATGGCTTTGAAC 123
QY 61 TCAGGGTCACCAACAGCTATTTGACCTTATGTAACCAAGTGAACCAACCGGAAAC 120
DB 124 TCAGGGTCACCAACAGCTATTTGACCTTATGTAACCAAGTGAACCAACCGGAAAC 183
QY 121 CCTATCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 180
DB 184 CCTATCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 243
QY 181 TACCGTCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 240
DB 244 TACCGTCCCGACAGCCCACTGTGTCCCACTGTCTACAGAGTGATCCGGCTCACTAC 303
QY 241 GTCTGACGAGCCCAATCCCATCC 267
DB 304 GTCTGACGAGCCCAATCCCATCC 330

RESULT 4
AF270487 1740 bp mRNA linear PRI 04-APR-2001
LOCUS Homo sapiens androgen-regulated serine protease TMPRSS2 precursor
DEFINITION (TMPRSS2) mRNA, complete cds.
ACCESSION AF270487
VERSION AF270487.1 GI:13540003
KEYWORDS

SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1740).
TITLE Afaar, D.B., Vivanco, I., Hubert, R.S., Kuo, J., Chen, E., Saffran, D.C., Raitano, A.B. and Jakobovits, A.
JOURNAL Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia
MEDLINE Cancer Res. 61 (4), 1686-1692 (2001)
PUBMED 21138112
11245484
2 (bases 1 to 1740)
REFERENCE Mitchell, S.C., Hubert, R.S. and Afaar, D.B.H.
AUTHORS Direct Submission
JOURNAL Submitted (19-MAY-2000) Urogenesis, Inc., 1701 Colorado Ave., Santa Monica, CA 90404, USA
FEATURES
source
1. 1740
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 1740
/gene="TMPRSS2"
114. 1592
/gene="TMPRSS2"
/note="expressed in prostate and prostate cancer; differs from product encoded by the sequence presented in GenBank
Accession Number U75329"
/codon_start=1
/product="androgen-regulated serine protease TMPRSS2 precursor"
/protein_id="AAK29280.1"
/db_xref="GI:13540004"
/translation="MALNSGSPALGIPYEMNGYOPENYPAPQTVPTVVEVHPAY
YSPVQVAPRYLTQASNPVCTOPKSPGTVCTSKTKALCTTLTGLFVLAALAA
GLMTFMGSKNSGIECDSSGTCINPSNMCDSVSHCPGSDENRCVRLYGGPFIQY
YSORSHVPCODNNENYGRACMDMGYKNNFVSQCIYVDSGTSMTKNTSGN
VDIYKLYSDACSSKAVYSLACIACGVNLNSRSGRTVGBSALPAMPQVSLHVO
NVHVGSGSLITPEWITTAHCEKELPNNMWHMTAFAGILRQSFMYGAYOVEKVI
ENYDSKTKNDIALMKDQPLTFNDLVKVCLEPNPGLMQLPGLCWIISMGATEEGK
TSEVNAALAVLIEFORCNSRYVDLITPAMICAPGLQANDVCGDSDGGLVYSKN
NTMTLIGDTSWCSGCAKAYRPVYGVVMTDITWIRQMRADG"
879. 1589
/gene="TMPRSS2"
/product="androgen-regulated serine protease TMPRSS2"
BASE COUNT 434 a 459 c 474 g 373 t
ORIGIN
Query Match 100.0%; Score 267; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 60
Db 66 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 125
Qy 61 TAGAGTCAACCAAGCTATGACCTTATGAAACCATGATGATCAACCGGAAAC 120
Db 126 TCAGGATCAACCAAGCTATGACCTTATGAAACCATGATGATCAACCGGAAAC 185
Qy 121 CCTATCCGCGACAGCCGCTGTGCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 180
Db 186 CCTATCCGCGACAGCCGCTGTGCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 245
Qy 181 TACCGGTCCCGGTGCGCCAGTACGCGGAGGTCTGACGAGGCTTCAACCCCGTC 240
Db 246 TACCGGTCCCGGTGCGCCAGTACGCGGAGGTCTGACGAGGCTTCAACCCCGTC 305
Qy 241 GTCTGACGACGACCAATCCCATCC 267
Db 306 GTCTGACGACGACCAATCCCATCC 332

RESULT 5
AX201121
LOCUS 2479 bp
DEFINITION Sequence 751 from Patent WO0151633.
ACCESSION AX201121
VERSION AX201121.1 GI:15390895
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2479)
AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W., Stolk, J.A., Skeiky, Y.A., Wang, A. and Meagher, M.J.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0151633-A 751 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 2479
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 578 a 650 c 677 g 574 t
ORIGIN
Query Match 100.0%; Score 267; DB 6; Length 2479;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 60
Db 9 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 66
Qy 61 TCAGGATCAACCAAGCTATGACCTTATGAAACCATGATGATCAACCGGAAAC 120
Db 69 TCAGGATCAACCAAGCTATGACCTTATGAAACCATGATGATCAACCGGAAAC 128
Qy 121 CCTATCCGCGACAGCCGCTGTGCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 180
Db 129 CCTATCCGCGACAGCCGCTGTGCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 188
Qy 181 TACCGGTCCCGGTGCGCCAGTACGCGGAGGTCTGACGAGGCTTCAACCCCGTC 240
Db 189 TACCGGTCCCGGTGCGCCAGTACGCGGAGGTCTGACGAGGCTTCAACCCCGTC 248
Qy 241 GTCTGACGACGACCAATCCCATCC 267
Db 249 GTCTGACGACGACCAATCCCATCC 275
RESULT 6
AX207965
LOCUS 2479 bp
DEFINITION Sequence 69 from Patent WO0157194.
ACCESSION AX207965
VERSION AX207965.1 GI:15422561
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2479)
AUTHORS Madison, E.L., Ong, S.O. and Yeh, J.C.
TITLE Nucleic acid molecules encoding transmembrane serine proteases, the encoded proteins and methods based thereon
JOURNAL Patent: WO 0157194-A 69 09-AUG-2001;
CORVUS INTERNATIONAL, INC. (US)
FEATURES
source
1. 2479
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
57. 1535

/note="Nucleotide sequence encoding human serine protease"
 /codon_start=1
 /protein_id="CAC60388.1"
 /db_xref="GI:15422562"
 /translation="MALNSGSPPAIGPYENHGYQENPYPAQPTVVPVYVHPDAQY
 VSPVPOYAPRVLTQASNPVCTQPKSPGCTVCTSKTKALCTITLTGFLVGAALAA
 GLTKFMGSKCSNGIECDSSGTCINPSCGVSHCPGEGDENRCVRLYGFNFILOM
 YSQRKSWHPVQDDWNNENYGRACDWMYKKNPFSSQGIYDSDSTSPMKLNTSAGN
 VDIYKLYHSDSSKAVVLRCLACGVNLNSRSGRIYGSFALPGAMPQVSLHVO
 NVHVGCGSIIITPEMIVTAHCEKPLNPMHTAFAGILROSEFMFYGAQYQVQKYSH
 PNYDSKTKNDIALMKLOKPLTENDLVKVCILPNEGMLQBPOLCMLISGMGATERRKX
 TSEVUNAAKVLILTORCNSRYVDNLITPMICAGFLQGNVDSQGDGSGPLVTSNN
 NIMWLIGTSMGSCAKAYRPGVGNVWFIDMIVYRKANG"

BASE COUNT 578 a 650 c 677 g 574 t

ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 6.2e-63;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATCTGATGCTGTGATGAACAAGATGCTTGAAC 60
 DB 9 GAACATTCCAGATACCTATCTGATGCTGTGATGAACAAGATGCTTGAAC 68
 QY 61 TCAGGGTCAACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 120
 DB 69 TCAGGGTCAACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 128
 QY 121 CCTATCCCGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCCGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 188
 QY 181 TACCGTCCCGGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 240
 DB 189 TACCGTCCCGGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 248
 QY 241 GTCTGACGACGCCCAATCCCATCC 267
 DB 249 GTCTGACGACGCCCAATCCCATCC 275

RESULT 7
 LOCUS AX267920 2479 bp DNA linear PAT 26-OCT-2001
 DEFINITION Sequence 894 from Patent WO0173032.
 ACCESSION AX267920
 VERSION AX267920.1 GI:16516526
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 AUTHORS 1
 Ku,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
 Kales,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
 Vedick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepner,W.T.
 and Henderson,R.A.
 TITLE Compositions and methods for the therapy and diagnosis of prostate
 cancer
 JOURNAL Patent: WO 0173032-A 894 04-OCT-2001;
 FEATURES
 SOURCE location/Qualifiers
 1..2479
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 578 a 650 c 677 g 574 t

Query Match 100.0%; Score 267; DB 6; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 6.2e-63;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATCTGATGCTGTGATGAACAAGATGCTTGAAC 60

DB 9 GAACATTCCAGATACCTATCTGATGCTGTGATGAACAAGATGCTTGAAC 68
 QY 61 TCAGGGTCAACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 120
 DB 69 TCAGGGTCAACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 128
 QY 121 CCTATCCCGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCCGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 188
 QY 181 TACCGTCCCGGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 240
 DB 189 TACCGTCCCGGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 248
 QY 241 GTCTGACGACGCCCAATCCCATCC 267
 DB 249 GTCTGACGACGCCCAATCCCATCC 275

RESULT 8
 LOCUS AX306771 2479 bp DNA linear PAT 11-DEC-2001
 DEFINITION Sequence 4 from Patent WO0018961.
 ACCESSION AX306771
 VERSION AX306771.1 GI:17645937
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 AUTHORS 1
 Macbeth,K.J. and Shyjan,A.W.
 TITLE Expression analysis of specific nucleic acids and polypeptides
 useful in the diagnosis and treatment of prostate cancer
 JOURNAL Patent: WO 0018961-A 4 06-APR-2000;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 SOURCE location/Qualifiers
 1..2479
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

CDS

/note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD19059.1"
 /db_xref="GI:17645938"
 /translation="MALNSGSPPAIGPYENHGYQENPYPAQPTVVPVYVHPDAQY
 YSPVPOYAPRVLTQASNPVCTQPKSPGCTVCTSKTKALCTITLTGFLVGAALAA
 GLTKFMGSKCSNGIECDSSGTCINPSCGVSHCPGEGDENRCVRLYGFNFILOM
 YSQRKSWHPVQDDWNNENYGRACDWMYKKNPFSSQGIYDSDSTSPMKLNTSAGN
 VDIYKLYHSDSSKAVVLRCLACGVNLNSRSGRIYGSFALPGAMPQVSLHVO
 NVHVGCGSIIITPEMIVTAHCEKPLNPMHTAFAGILROSEFMFYGAQYQVQKYSH
 PNYDSKTKNDIALMKLOKPLTENDLVKVCILPNEGMLQBPOLCMLISGMGATERRKX
 TSEVUNAAKVLILTORCNSRYVDNLITPMICAGFLQGNVDSQGDGSGPLVTSNN
 NIMWLIGTSMGSCAKAYRPGVGNVWFIDMIVYRKANG"

BASE COUNT 578 a 650 c 677 g 574 t

ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 6.2e-63;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATACCTATCTGATGCTGTGATGAACAAGATGCTTGAAC 60
 DB 9 GAACATTCCAGATACCTATCTGATGCTGTGATGAACAAGATGCTTGAAC 68
 QY 61 TCAGGGTCAACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 120
 DB 69 TCAGGGTCAACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 128
 QY 121 CCTATCCCGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCCGACAGCCACTGTGCTCCCACTAGAGGTGATCCGGCTCAGTAC 188

QY 181 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGAGCCCAATCCCATCC 267
 Db 249 GTCTGACGAGCCCAATCCCATCC 275
 RESULT 9
 LOCUS AX395318 2479 bp DNA linear PAT 18-MAY-2002
 DEFINITION Sequence 3 from Patent WO0204953.
 ACCESSION AX395318
 VERSION AX395318.1 GI:21066314
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 AUTHORS Saferian, D., Rattano, A.B., Hubert, R.S., Jakobovits, A., Fairs, M. and
 Challita-Eid, P.M.
 TITLE Novel tumor antigen useful in diagnosis and therapy of bladder,
 ovary, lung and kidney cancers
 JOURNAL Patent: WO 0204953-A 3 17-JAN-2002;
 Agensys, Inc. (US)
 FEATURES
 source 1..2479
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 578 a 650 c 677 g 574 t
 ORIGIN
 Query Match 100.0%; Score 267; DB 6; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 6.2e-63;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGAGTACCTATCTAGTGTGATTAACAGCAAGTGGCTTTGAAC 60
 Db 9 GAACATTCGAGTACCTATCTAGTGTGATTAACAGCAAGTGGCTTTGAAC 68
 QY 61 TCAGGGTACACAGCTATTTGACCTTATGAAAACATGATACCAACCGGAAAC 120
 Db 69 TCAGGGTACACAGCTATTTGACCTTATGAAAACATGATACCAACCGGAAAC 128
 QY 121 CCTATCCGCGACAGCCCACTGTGTCTGACGAGTGCATCCGGCTCAGTAC 180
 Db 129 CCTATCCGCGACAGCCCACTGTGTCTGACGAGTGCATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGAGCCCAATCCCATCC 267
 Db 249 GTCTGACGAGCCCAATCCCATCC 275
 RESULT 10
 LOCUS HSU75329 2479 bp mRNA linear PRI 10-OCT-1997
 DEFINITION Human serine protease mRNA, complete cds.
 ACCESSION U75329
 VERSION U75329.1 GI:2507612
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2479)
 Paoloni-Giacobino, A., Chen, H., Peltsch, M.C., Rosier, C. and

TITLE Antarakis, S.B.
 JOURNAL Cloning of the TMPRSS2 gene, which encodes a novel serine protease
 MEDLINE with transmembrane, LDIRA, and SRCR domains and maps to 21q22.3
 97468144
 PUBMED Genomics 44 (3), 309-320 (1997)
 REFERENCE 2 (bases 1 to 2479)
 9325052
 AUTHORS Paoloni-Giacobino, A., Chen, H. and Antarakis, S.B.
 TITLE Direct Submission
 JOURNAL Submitted (17-OCT-1996) Medical Genetics, University of Geneva
 Medical School, 1 Michel-Servet Street, Geneva 1211, Switzerland
 FEATURES
 source 1..2479
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 57..1535
 /codon_start=1
 /product="serine protease"
 /protein_id="AAC51784.1"
 /db_xref="GI:2507613"
 /translation="MALNSGSPPAIGPYENHGYOPENPYPAQPTVPTVYEVHPAY
 YPSVPQYAPRVLTQASNPVCTOPKSPSGCTGSKTKKALCTTITGLTGLVGAALA
 GLMTKMGSKCSNGIECDSSGTCINPSNCDGVSHCPGSEDNRCVRLKYPNFIOM
 YSQRKSWHPVCDQDMWENTGRACRDMGKNNFYSSQIVDSGTSFPMKNTLSAGN
 VDIYKLKTHSDACSRAVSLRCLACGVNINSRORIVGESALPGAMPQVSLHYO
 NNAVCGSIIITPEMIIVTAACHVEKPLNNPWTAPAGILRDSFMYGAGYQVOKVSH
 PNYDSKTQNDIALMKLQKPLFENDLVKPCLEPNPMLOPBLCMFISGMGATBERGK
 TSEVLNAAKVLIETORCNSRYVDLITPAMICAFPLQGNVDSGDSGGPLVTSNN
 NIMWLIGDTSWSSGCAKAPRGVGNVMTFTDITIRKANG"
 BASE COUNT 578 a 650 c 677 g 574 t
 ORIGIN
 Query Match 100.0%; Score 267; DB 9; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 6.2e-63;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGAGTACCTATCTAGTGTGATTAACAGCAAGTGGCTTTGAAC 60
 Db 9 GAACATTCGAGTACCTATCTAGTGTGATTAACAGCAAGTGGCTTTGAAC 68
 QY 61 TCAGGGTACACAGCTATTTGACCTTATGAAAACATGATACCAACCGGAAAC 120
 Db 69 TCAGGGTACACAGCTATTTGACCTTATGAAAACATGATACCAACCGGAAAC 128
 QY 121 CCTATCCGCGACAGCCCACTGTGTCTGACGAGTGCATCCGGCTCAGTAC 180
 Db 129 CCTATCCGCGACAGCCCACTGTGTCTGACGAGTGCATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGAGCCCAATCCCATCC 267
 Db 249 GTCTGACGAGCCCAATCCCATCC 275
 RESULT 11
 LOCUS AF329454 3226 bp mRNA linear PRI 16-MAY-2001
 DEFINITION Homo sapiens epithelialin (TMPRSS2) mRNA, complete cds.
 ACCESSION AF329454
 VERSION AF329454.1 GI:14091027
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3226)
 Jacquinet, E., Rao, N.V., Rao, G.V., Zhengming, W., Albertine, K.H. and
 Hoidal, J.R.

TITLE Cloning and characterization of the cDNA and gene for human epithelialin

JOURNAL Eur. J. Biochem. 268 (9), 2687-2699 (2001)

MEDLINE 2123025

PUBMED 11322890

REFERENCE 2 (bases 1 to 3226)

AUTHORS Jacquinet, E., Rao, N.V., Rao, G.V., Wang, Z., Albertine, K.H. and Hoidal, J.R.

TITLE Direct Submission

JOURNAL Submitted (15-DEC-2000) Pulmonary Medicine, University of Utah, 50N. Medical Dr., Salt Lake City, UT 84132, USA

FEATURES

source

1..3226

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

1..3226

/gene="TMPRSS2"

128..1606

/gene="TMPRSS2"

/note="proteinase; PRSS10"

/codon_start=1

/product="epithelialin"

/protein_id="AAK53559.1"

/db_xref="GI:14091028"

/translation="MALNSGSPPAIGPYENHGYOPENPYAPPTVTVTVYEVHAPQYPSVPVQAPRVLTOASNPVCTOPKSPSGVCTKTKALCTITLGLTFLVGAALAAGLMFGMSKCSNGSIEDSGGCTINPNCNCGGGEDENCTVLYGNPILCMYSSORSHWPCDDPMENYGRRAACKMGYKNNPFSGGIYDDSDSTFPMCLNSAGVDIYKLIHSDSCSKAVSLACIACGVNLSSRSRLVGGSSALPGAMPQVSLHVO NVHVGSSIIITPEWLVTAHCEKPLNPMWTAGILROSFMFYAGYVEKVI SHPNYSKTNNDIALMKLQKPLFNDLVPLCPNGMMLQEBOLCMISGMATEKSKTSEVLNAKAVLLIETORCNSRYVDNLITPAMI CAGFLQGNVDSQQDSDGSPVLTSKN NIMWLIIGTSMWGCAKAYRPGVGNVWVFTDWIRQRADG"

BASE COUNT 782 a 808 c 857 g 779 t

ORIGIN

Query Match 100.0%; Score 267; DB 9; Length 3226;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGACGAAGTGGCTTGAAC 60
Db 80 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGACGAAGTGGCTTGAAC 139

Qy 61 TCAGGGTACCAACGCTATGGAACCTTACTATGAACCAATGATACCAACCGAATAAC 120
Db 140 TCAGGGTACCAACGCTATGGAACCTTACTATGAACCAATGATACCAACCGAATAAC 199

Qy 121 CCTATCCCGACAGCCCACTGTGCTCCCACTGTCTACGAGGTGATCCGGCTCAGTAC 180
Db 200 CCTATCCCGACAGCCCACTGTGCTCCCACTGTCTACGAGGTGATCCGGCTCAGTAC 259

Qy 181 TACCGGTCCCCGGTCCCGCAGTACGCCGCCGAGGGTCTGACCAAGCTTCAACCCCGTC 240
Db 260 TACCGGTCCCCGGTCCCGCAGTACGCCGCCGAGGGTCTGACCAAGCTTCAACCCCGTC 319

Qy 241 GTCTGACGACGACCAATCCCATCC 267
Db 320 GTCTGACGACGACCAATCCCATCC 346

RESULT 12

AX201156 3245 bp DNA linear PAT 29-AUG-2001

LOCUS AX201156

DEFINITION Sequence 786 from Patent WO0151633.

ACCESSION AX201156

VERSION AX201156.1 GI:15390911

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 3245)

AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W., Stolk, J.A., Skelky, Y.A., Wang, A. and Mesinger, M.J.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0151633-A 786 19-JUL-2001;

FEATURES

source

1..3245

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 774 a 818 c 870 g 783 t

ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 3245;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGACGAAGTGGCTTGAAC 60
Db 97 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGACGAAGTGGCTTGAAC 156

Qy 61 TCAGGGTACCAACGCTATGGAACCTTACTATGAACCAATGATACCAACCGAATAAC 120
Db 157 TCAGGGTACCAACGCTATGGAACCTTACTATGAACCAATGATACCAACCGAATAAC 216

Qy 121 CCTATCCCGACAGCCCACTGTGCTCCCACTGTCTACGAGGTGATCCGGCTCAGTAC 180
Db 217 CCTATCCCGACAGCCCACTGTGCTCCCACTGTCTACGAGGTGATCCGGCTCAGTAC 276

Qy 181 TACCGGTCCCCGGTCCCGCAGTACGCCGCCGAGGGTCTGACGACGAGCTTCAACCCCGTC 240
Db 277 TACCGGTCCCCGGTCCCGCAGTACGCCGCCGAGGGTCTGACGACGAGCTTCAACCCCGTC 336

Qy 241 GTCTGACGACGACCAATCCCATCC 267
Db 337 GTCTGACGACGACCAATCCCATCC 363

RESULT 13

AX267955 3245 bp DNA linear PAT 26-OCT-2001

LOCUS AX267955

DEFINITION Sequence 929 from Patent WO0173032.

ACCESSION AX267955

VERSION AX267955.1 GI:16516540

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1

AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H., Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skelky, Y.A., Hepner, W.T. and Henderson, R.A.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0173032-A 929 04-OCT-2001;

FEATURES

source

1..3245

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 774 a 818 c 870 g 783 t

ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 3245;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGACGAAGTGGCTTGAAC 60


```

Db 97 GAACATTCAGATACCTTACTTACTGATGCTGTGATTAACAGAGATGCTTGAAC 156
Qy 61 TCAGGGTCAACCAAGCTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 120
Db 157 TCAGGGTCAACCAAGCTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAC 216
Qy 121 CCCTATCCCGACAGCCCACTGTGTGCTCCCACTGTCTAAGAGTGCATCCGCTCAGTAC 180
Db 217 CCCTATCCCGACAGCCCACTGTGTGCTCCCACTGTCTAAGAGTGCATCCGCTCAGTAC 276
Qy 181 TACCGTCCCCCGTGGCCCACTGAGCCCGGAGGCTCTGACGCGAGCTTCCAAACCCGCTC 240
Db 277 TACCGTCCCCCGTGGCCCACTGAGCCCGGAGGCTCTGACGCGAGCTTCCAAACCCGCTC 336
Qy 241 GTCTGACAGCAGCCCAATCCCATCC 267
Db 337 GTCTGACAGCAGCCCAATCCCATCC 363

RESULT 14
LOCUS AX041973 3966 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 3 from Patent W00065067.
ACCESSION AX041973
VERSION AX041973.1 GI:11340738
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 3966)
AUTHORS Nelson,P.S., Hood,L. and Lin,B.
TITLE Prostate-specific polynucleotides, polypeptides and their methods
of use
JOURNAL Patent: WO 0065067-A 3 02-NOV-2000;
The University of Washington (US)
FEATURES
source location/Qualifiers
1..3966
/organism="Homo sapiens"
/db_xref="taxon:9606"
57..1535
/feature="unlabeled protein product"
/codon_start=1
/protein_id="CAC17064.1"
/db_xref="GI:11340738"
/translation="MALNSGSPPAIGPTPTENHGYOPENPYPAQPTVPTVYEVHAPQY
VSPVPIAPKVLTAQNSNVCTOPKSPGCTVSTKTKALCTTLTGFLVGAALAA
GLLMKFMGKCSNSGIECDSSGTCINPSCWDGSHCPGSDENRCRLTGFLVGAALAA
VSSORSMHPVCODDWNENYGRACADMDGKNNFVSSGIVDSDGTSFMKNTSAGN
VDIYKLYSDACSSKAVVSLRGLACGVALNSRORIRIVGSGALPGAMPVQVSLVQ
NVHVGGSITTPEMITVTAHCTBKPLNNHMTAFAIGIIRQSFMTYAGIYQVKVISH
PMDKSTKRNDAALAKLQKPLFNDLVKVCLEPNPQMLQPEQLCWSMGALTEBKGK
TSEVLAAPVLLIETQRCNSRYVDNLITPAMICAGFLQGNVDS CGSDSGPLVTSNN
NIMWLIQDTSWGSCKAYRPAVYGVNMTFTWITVYOMKANG"
BASE COUNT 992 a 998 c 988 g 984 t 4 others
ORIGIN
Query Match 100.0%; Score 267; DB 6; Length 3966;
Best Local Similarity 100.0%; Pred. No. 6.3e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 181 TACCGTCCCCCGTGGCCCACTGAGCCCGGAGGCTCTGACGCGAGCTTCCAAACCCGCTC 240
Db 189 TACCGTCCCCCGTGGCCCACTGAGCCCGGAGGCTCTGACGCGAGCTTCCAAACCCGCTC 248
Qy 241 GTCTGACAGCAGCCCAATCCCATCC 267
Db 249 GTCTGACAGCAGCCCAATCCCATCC 275

RESULT 15
LOCUS AX201158 1476 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 788 from Patent W00151633.
ACCESSION AX201158
VERSION AX201158.1 GI:15390913
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1476)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
Stolk,J.A., Skelky,Y.A., Wang,A. and Meagher,M.U.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0151633-A 788 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..1476
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 361 a 405 c 402 g 308 t
ORIGIN
Query Match 82.0%; Score 219; DB 6; Length 1476;
Best Local Similarity 100.0%; Pred. No. 1e-49;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 49 ATGGCTTTGAAGTCAAGGGTCAACCAAGCTATTGACCTTACTATGAAAACCATGATAC 108
Db 1 ATGGCTTTGAAGTCAAGGGTCAACCAAGCTATTGACCTTACTATGAAAACCATGATAC 60
Qy 109 CAACCGGAAAACCCCTATCCGACAGCCCACTGTGTGCCACTGTCTAAGAGTGCAT 168
Db 61 CAACCGGAAAACCCCTATCCGACAGCCCACTGTGTGCCACTGTCTAAGAGTGCAT 120
Qy 169 CCGGCTCAGACTACCCGCTCCCGGCTCCCGGCTCCAGTACGCCCGGAGGTCTGACGAGGCT 228
Db 121 CCGGCTCAGACTACCCGCTCCCGGCTCCCGGCTCCAGTACGCCCGGAGGTCTGACGAGGCT 180
Qy 229 TCCAACCCCGTGTGTGACAGCAGCCCAATCCCATCC 267
Db 181 TCCAACCCCGTGTGTGACAGCAGCCCAATCCCATCC 219

```

Search completed: January 31, 2003, 07:50:50
 Job time : 3222 secs